



IPW

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re U.S. Patent Application of)
HIGAKI et al.)
Application Number: 10/768,439)
Filed: February 2, 2004)
For: STORAGE SYSTEM AND STORAGE CONTROL)
DEVICE)
Attorney Docket No. HITA.0508)

**Honorable Assistant Commissioner
for Patents
Washington, D.C. 20231**

LETTER

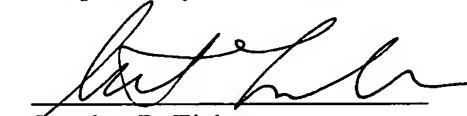
Sir:

The below-identified communications are submitted in the above-captioned application or proceeding:

(X) Priority Documents (1)	() Assignment Document
(X) Request for Priority	() Petition under 37 C.F.R. 1.47(a)
() Response to Missing Parts w/ signed Declaration	() Check for \$130.00

The Commissioner is hereby authorized to charge payment of any fees associated with this communication, including fees under 37 C.F.R. § 1.16 and 1.17 or credit any overpayment to Deposit Account Number 08-1480. A duplicate copy of this sheet is attached.

Respectfully submitted,



Stanley P. Fisher
Registration Number 24,344

Juan Carlos A. Marquez
Registration Number 34,072

REED SMITH LLP
3110 Fairview Park Drive
Suite 1400
Falls Church, Virginia 22042
(703) 641-4200
June 29, 2004



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re U.S. Patent Application of)
HIGAKI et al.)
Application Number: 10/768,439)
Filed: February 2, 2004)
For: STORAGE SYSTEM AND STORAGE CONTROL)
DEVICE)
Attorney Docket No. HITA.0508)

**Honorable Assistant Commissioner
for Patents
Washington, D.C. 20231**

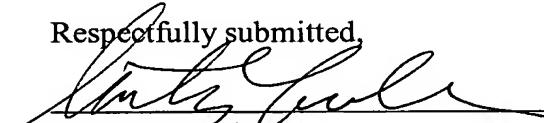
**REQUEST FOR PRIORITY
UNDER 35 U.S.C. § 119
AND THE INTERNATIONAL CONVENTION**

Sir:

In the matter of the above-captioned application for a United States patent, notice is hereby given that the Applicant claims the priority date of September 16, 2003 the filing date of the corresponding Japanese patent application 2003-323120.

A certified copy of Japanese patent application 2003-323120 is being submitted herewith. Acknowledgment of receipt of the certified copy is respectfully requested in due course.

Respectfully submitted,



Stanley P. Fisher
Registration Number 24,344

REED SMITH LLP
3110 Fairview Park Drive
Suite 1400
Falls Church, Virginia 22042
(703) 641-4200
June 29, 2004

Juan Carlos A. Marquez
Registration Number 34,072

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 9月16日
Date of Application:

出願番号 特願2003-323120
Application Number:

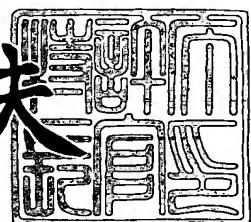
[ST. 10/C] : [JP2003-323120]

出願人 株式会社日立製作所
Applicant(s):

2004年 2月19日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



【書類名】 特許願
【整理番号】 340300757
【あて先】 特許庁長官殿
【国際特許分類】 G06F 03/06
【発明者】
 【住所又は居所】 神奈川県小田原市中里 322 番 2 号 株式会社日立製作所 R A I
 D システム事業部内
 【氏名】 檜垣 誠一
【発明者】
 【住所又は居所】 神奈川県小田原市中里 322 番 2 号 株式会社日立製作所 R A I
 D システム事業部内
 【氏名】 島田 朗伸
【発明者】
 【住所又は居所】 神奈川県小田原市中里 322 番 2 号 株式会社日立製作所 R A I
 D システム事業部内
 【氏名】 岡見 吉規
【発明者】
 【住所又は居所】 神奈川県小田原市中里 322 番 2 号 株式会社日立製作所 R A I
 D システム事業部内
 【氏名】 中野 俊夫
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社日立製作所
【代理人】
 【識別番号】 100095371
 【弁理士】
 【氏名又は名称】 上村 輝之
【選任した代理人】
 【識別番号】 100089277
 【弁理士】
 【氏名又は名称】 宮川 長夫
【選任した代理人】
 【識別番号】 100104891
 【弁理士】
 【氏名又は名称】 中村 猛
【手数料の表示】
 【予納台帳番号】 043557
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1
 【包括委任状番号】 0110323

【書類名】特許請求の範囲**【請求項 1】**

第1の記憶制御装置と第2の記憶制御装置とを通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶システムであって、

前記第1の記憶制御装置は、

前記上位装置から受信した第1のリクエストに係る所定のデータ処理を前記第2の記憶制御装置が実行可能か否かを判断し、前記第2の記憶制御装置が実行可能と判断した場合は、前記第1のリクエストに対応する第2のリクエストを生成して前記第2の記憶制御装置に送信させる第1の制御手段を有し、

前記第2の記憶制御装置は、

前記第1の記憶制御装置から受信した前記第2のリクエストに基づいて、前記所定のデータ処理を行う第2の制御手段を有する、

ことを特徴とする記憶システム。

【請求項 2】

前記第1の記憶制御装置は、前記第2の記憶制御装置が管理する第2の記憶領域を、自己の管理する第1の記憶領域として前記上位装置に仮想的に提供するものであり、

前記第1のリクエストは、前記第1の記憶領域に関するデータ処理を要求するものである請求項1に記載の記憶システム。

【請求項 3】

前記第1の記憶制御装置は、前記第1の記憶領域と前記第2の記憶領域との対応関係を示す記憶領域対応情報を保持し、該記憶領域対応情報に基づいて前記上位装置に第1の記憶領域を仮想的に提供するものであり、

前記第1の制御手段は、前記記憶領域対応情報に基づいて、前記第1のリクエストによる前記第1の記憶領域を対象とするデータ処理を実行可能である請求項2に記載の記憶システム。

【請求項 4】

前記第2のリクエストは、前記第1のリクエストと同様のデータ構造を有する請求項1に記載の記憶システム。

【請求項 5】

前記第1の制御手段は、前記第2のリクエストを前記第2の記憶制御装置に送信する前に、前記第2のリクエストに係る所定のデータ処理を前記第2の記憶制御装置が実行可能か否かを確認する請求項1に記載の記憶システム。

【請求項 6】

前記第1の記憶制御装置は、前記第2の記憶制御装置で実行可能なデータ処理機能を示す機能管理情報を保持し、

前記第1の制御手段は、前記機能管理情報に基づいて、前記第2のリクエストに係る所定のデータ処理を前記第2の記憶制御装置で実行可能か否かを判断する請求項1に記載の記憶システム。

【請求項 7】

前記機能管理情報は、記憶システムの構成を定義する際に、手動で又は自動的に生成されるものである請求項6に記載の記憶システム。

【請求項 8】

前記第1の記憶制御装置及び前記第2の記憶制御装置の両方に通信可能に接続されたバックアップ装置を備え、

前記第1の制御手段は、

前記第1のリクエストに係るデータ処理が前記第1の記憶領域に記憶されている情報を前記バックアップ装置に転送して記憶させるバックアップ処理である場合、前記第2の記憶制御装置が前記バックアップ処理を実行可能か否かを判断し、前記第2の記憶制御装置が前記バックアップ処理を実行可能と判断した場合は、前記第1のリクエストに含まれる前記第1の記憶領域におけるアドレスを前記第2の記憶領域におけるアドレスに変換する

ことにより、前記第2のリクエストを生成して前記第2の記憶制御装置に送信するものであります。

前記第2の制御手段は、

前記第2のリクエストに基づいて、前記第2の記憶領域に記憶されている情報を前記バックアップ装置に転送して記憶させるものである請求項2に記載の記憶システム。

【請求項9】

前記第1の記憶制御装置は、第1の記憶領域とペアを構成する第1のセカンダリ記憶領域をさらに仮想的に提供するものであり、

前記第2の記憶制御装置は、前記第2の記憶領域とペアを構成する第2のセカンダリ記憶領域をさらに有し、

前記第1の制御手段は、

前記第1のリクエストに係る所定のデータ処理が前記第1の記憶領域に記憶されている情報を前記第1のセカンダリ記憶領域にコピーする内部コピー処理である場合、前記第2の記憶制御装置が前記内部コピー処理を実行可能か否かを判断し、前記第2の記憶制御装置が前記内部コピー処理を実行可能と判断した場合は、前記第1のリクエストに含まれる前記第1の記憶領域におけるアドレスを前記第2の記憶領域におけるアドレスに変換することにより、前記第2のリクエストを生成して前記第2の記憶制御装置に送信するものであります。

前記第2の制御手段は、

前記第2のリクエストに基づいて、前記第2の記憶領域に記憶されている情報を前記第2のセカンダリ記憶領域にコピーさせるものである請求項2に記載の記憶システム。

【請求項10】

前記第1の記憶制御装置及び前記第2の記憶制御装置が設けられる正サイトと対をなし、正サイトと通信可能に接続された副サイトを有し、

前記副サイトは、他の第1の記憶制御装置及び他の第2の記憶制御装置を備え、前記他の第1の記憶制御装置は、前記他の第2の記憶制御装置が管理する他の第2の記憶領域を、自己の管理する他の第1の記憶領域として仮想的に提供するものであり、

前記正サイトの第1の制御手段は、

前記第1のリクエストに係る所定のデータ処理が前記第1の記憶領域に記憶されている情報を前記副サイトの他の第1の記憶領域にコピーする外部コピー処理である場合、前記第2の記憶制御装置及び前記他の第2の記憶制御装置の両方が前記外部コピー処理を実行可能か否かを判断し、前記各第2の記憶制御装置がそれぞれ前記外部コピー処理を実行可能と判断した場合は、前記第1のリクエストに対応する第2のリクエストを生成して前記第2の記憶制御装置に送信させ、

前記第2の制御手段は、前記第2のリクエストに基づいて、前記第2の記憶領域に記憶されている情報を前記他の第2の記憶領域にコピーされることにより前記外部コピー処理を実行する請求項2に記載の記憶システム。

【請求項11】

前記他の第1の記憶制御装置は、前記他の第2の記憶制御装置で実行可能なデータ処理機能を示す他の機能管理情報を保持しており、

前記第1の制御手段は、前記第2のリクエストを前記第2の記憶制御装置に送信する前に、前記他の第1の記憶制御装置に問い合わせることにより、前記他の第2の記憶制御装置が前記外部コピー処理を実行可能か否かを判断するものである請求項10に記載の記憶システム。

【請求項12】

前記第1の記憶制御装置は、前記第1の記憶領域と前記第2の記憶領域との対応関係を示す記憶領域対応情報を保持し、

前記他の第1の記憶制御装置は、前記他の第1の記憶領域と前記他の第2の記憶領域との対応関係を示す他の記憶領域対応情報を保持し、

前記第1の制御手段は、

前記第2のリクエストを前記第2の記憶制御装置に送信する場合に、前記各記憶領域対応情報を送信する請求項10に記載の記憶システム。

【請求項13】

前記第1の記憶制御装置は、前記外部コピー処理中に前記上位装置により更新された前記第1の記憶領域の情報に関する情報を保持する更新位置情報保持手段を備え、

前記第1の制御手段は、前記外部コピー処理が完了した場合に、前記更新位置情報保持手段に基づいて、前記第1の記憶領域において更新された情報を前記他の第1の記憶領域に記憶させるべく、前記第2のリクエストを生成して、前記第2の記憶制御装置から更新された情報を読み出して、前記読み出した情報を送信する請求項10に記載の記憶システム。

【請求項14】

第1の記憶制御装置と第2の記憶制御装置とを通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶システムを制御するための制御方法であって、

前記第1の記憶制御装置は、

前記上位装置から第1のリクエストを受信するステップと、

前記受信した第1のリクエストに係る所定のデータ処理を前記第2の記憶制御装置が実行可能か否かを判断するステップと、

前記第2の記憶制御装置が実行可能と判断した場合は、前記第1のリクエストに対応する第2のリクエストを生成するステップと、

前記生成された第2のリクエストを前記第2の記憶制御装置に送信させるステップとを実行し、

前記第2の記憶制御装置は、

前記第1の記憶制御装置から前記第2のリクエストを受信するステップと、

前記受信した第2のリクエストに基づいて、前記所定のデータ処理を行うステップとを実行する記憶システムの制御方法。

【請求項15】

第2の記憶制御装置及び上位装置と通信可能に接続され、前記上位装置からのリクエストに応じたデータ処理を行う記憶制御装置であって、

前記上位装置から第1のリクエストを受信する受信手段と、

前記受信した第1のリクエストに係わる所定のデータ処理を前記第2の記憶制御装置が実行可能か否かを判断する判断手段と、

前記第2の記憶制御装置が実行可能と判断した場合は、前記第1のリクエストに対応する第2のリクエストを生成するリクエスト手段と、

前記生成された第2のリクエストを前記第2の記憶制御装置に送信する送信手段と、を備えて構成される記憶制御装置。

【請求項16】

第2の記憶制御装置及び上位装置と通信可能に接続され、前記上位装置からのリクエストに応じたデータ処理を行う第1の記憶制御装置を制御するプログラムであって、

前記上位装置から受信した第1のリクエストに係る所定のデータ処理を前記第2の記憶制御装置が実行可能か否かを判断する機能と、

前記第2の記憶制御装置が実行可能と判断した場合は、前記第1のリクエストに対応する第2のリクエストを生成する機能と、

前記生成された第2のリクエストを前記第2の記憶制御装置に送信させる機能とを、前記第1の記憶制御装置のコンピュータ上に実現させるプログラム。

【書類名】明細書

【発明の名称】記憶システム及び記憶制御装置

【技術分野】

【0001】

本発明は、記憶システム及び記憶制御装置に関する。

【背景技術】

【0002】

例えば、データセンタ等のような大規模なデータを取り扱うデータベースシステムでは、ホストコンピュータとは別に構成された記憶システムを用いてデータを管理する。この記憶システムは、例えば、ディスクアレイ装置等から構成される。ディスクアレイ装置は、多数のディスク記憶装置をアレイ状に配設して構成されるもので、例えば、RAID (Redundant Array of Independent Inexpensive Disks) に基づいて構築されている。ディスク装置群が提供する物理的な記憶領域上には少なくとも1つ以上の論理ボリュームが形成され、この論理ボリュームがホストコンピュータ（より詳しくは、ホストコンピュータ上で稼働するデータベースプログラム）に提供される。ホストコンピュータは、所定のコマンドを送信することにより、論理ボリュームに対してデータの書き込み、読み出しを行うことができる。

【0003】

情報化社会の進展等につれて、データベースで管理すべきデータは、日々増大する。このため、より高性能、より大容量の記憶制御装置が求められており、この市場要求に応えるべく、新型の記憶制御装置が開発されている。新型の記憶制御装置を記憶システムを導入する方法としては、2つの方法が考えられる。その一つは、旧型の記憶制御装置と新型の記憶制御装置とを完全に入れ替え、全て新型の記憶制御装置から記憶システムを構成する方法である（特許文献1）。他の一つは、旧型の記憶制御装置からなる記憶システムに新型の記憶制御装置を新たに追加し、新旧の記憶制御装置を併存させる方法である。

【特許文献1】特表平10-508967号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

旧型の記憶制御装置から新型の記憶制御装置に完全に移行する場合（特許文献1）は、新型の記憶制御装置の機能、性能を利用することができるが、旧型の記憶制御装置を有効に利用することができない。他方、旧型の記憶制御装置と新型の記憶制御装置との併存を図る場合は、記憶システムを構成する記憶制御装置の数が増大し、新旧の記憶制御装置を管理し運用する手間が大きい。

【0005】

本発明は、上記の問題点に鑑みてなされたもので、本発明の目的の一つは、特定の記憶制御装置に負荷が集中するのを防止し、負荷を分散させることができる記憶システム及び記憶制御装置を提供することにある。

本発明の別の目的は、例えば、新旧の記憶制御装置のように異なる記憶制御装置を連携させて、記憶資源の論理的な集中化と負荷の分散とを両立可能な記憶システム及び記憶制御装置を提供することにある。

本発明の他の目的は、例えば、記憶システムの高性能化大容量化を実現しつつ、高機能高性能な記憶制御装置に処理負荷が集中するのを防止し、負荷を分散させることができるようとした記憶システム及び記憶制御装置を提供することにある。

本発明の更なる目的は、後述する実施の形態の記載から明らかになるであろう。

【課題を解決するための手段】

【0006】

上記課題を解決すべく、本発明に従う記憶システムは、第1の記憶制御装置と第2の記憶制御装置とを通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶システムであって、第1の記憶制御装置は、上位装置から受信した第1の

リクエストに係る所定のデータ処理を第2の記憶制御装置が実行可能か否かを判断し、第2の記憶制御装置が実行可能と判断した場合は、第1のリクエストに対応する第2のリクエストを生成して第2の記憶制御装置に送信させる第1の制御手段を有し、第2の記憶制御装置は、第1の記憶制御装置から受信した第2のリクエストに基づいて、所定のデータ処理を行う第2の制御手段を有する、ことを特徴とする。

【0007】

記憶制御装置としては、例えば、ディスクアレイ装置やファイバチャネルスイッチ等を挙げることができる。上位装置としては、例えば、パーソナルコンピュータ、メインフレーム等のコンピュータを挙げることができる。第1の記憶制御装置と第2の記憶制御装置とは、通信ネットワークを介して双方向通信可能に接続されており、第1の記憶制御装置と上位装置も通信ネットワークを介して双方向通信可能に接続されている。また、第2の記憶制御装置と上位装置との間も双方向通信可能に接続することができる。通信ネットワークとしては、例えば、LAN (Local Area Network)、SAN (Storage Area Network)、専用回線、インターネット等を挙げることができる。

【0008】

上位装置は、第1の記憶制御装置に第1のリクエストを送信する。このリクエストは、例えば、要求内容を特定するためのコマンドコード、対象とするデータのアドレスを特定するアドレス情報等を含んで構成される。第1の記憶制御装置の第1の制御手段は、第1のリクエストを受信すると、第1のリクエストで要求されている所定の処理を、第2の記憶制御装置で実行可能か否かを判断する。ここで、所定の処理としては、種々のデータ処理を挙げることができるが、例えば、単純なデータ入出力以外のサポート機能（附加的機能）を挙げができる。具体的には、例えば、データのバックアップ、ペアボリューム間のコピー、ミラーリング等のような、比較的処理負荷の大きな処理を所定の処理とすれば、大きな負荷を第1の記憶制御装置以外の装置に分散させることができる。

【0009】

第1の制御手段は、第2の記憶制御装置が所定の処理を実行できると判定した場合、第1のリクエストに対応する第2のリクエストを生成し、第2の記憶制御装置に送信する。第2のリクエストは、所定の処理を第2の記憶制御装置に行わせることを要求する情報である。第2の記憶制御装置が第2のリクエストを受信すると、第2の制御手段は、第2のリクエストに基づいて、要求された所定のデータ処理を実行する。このように、第1のリクエストによって第1の記憶制御装置に要求されたデータ処理を、第2のリクエストによって第2の記憶制御装置に代行して実行させることができる。

【0010】

従って、第1の記憶制御装置の処理負荷を第2の記憶制御装置に分散させて、第1の記憶制御装置に過大な負荷が集中するのを防止することができる。そして、第1の記憶制御装置の情報処理資源（CPU処理能力やメモリ容量等）を、上位装置へのサービス提供に割り当てることができる。第1の記憶制御装置が新型の記憶制御装置である場合、第1の記憶制御装置が有する高性能なサービスを上位装置に有効に提供することができ、記憶システム全体の効率を向上させることができる。

【0011】

本発明の一態様では、第1の記憶制御装置は、第2の記憶制御装置が管理する第2の記憶領域を、自己の管理する第1の記憶領域として上位装置に仮想的に提供し、第1のリクエストは、第1の記憶領域に関するデータ処理を要求するものである。

【0012】

第2の記憶制御装置は、例えば、ディスクドライブ等の記憶装置により提供される現実の記憶領域を備えている。この実際に存在する記憶領域を、本明細書では実記憶領域とも呼ぶ。第1の記憶制御装置は、第2の記憶制御装置の実記憶領域である第2の記憶領域を、あたかも自己の有する記憶領域のように見せかけて、上位装置に仮想的に提供する。従って、第1の記憶制御装置自身は、物理的な記憶領域を備えている必要はなく、マイクロコンピュータシステム等を有するインテリジェント化されたスイッティング機構（ファイバ

チャネルスイッチ等)を、第1の記憶制御装置として用いることができる。なお、第1の記憶制御装置は、物理的な記憶領域を備えたディスクアレイ装置(ディスクアレイサブシステム)等であってもよい。第1の記憶制御装置をディスクアレイ装置から構成した場合は、第2の記憶制御装置の第2の記憶領域を取り込む結果、実際に有する記憶容量よりも大きな記憶領域を上位装置に対して提供することができる。

【0013】

上位装置は、第1のリクエストを発行することにより、仮想的に提供された第1の記憶領域に対してデータ操作を要求するが、実際のデータは第2の記憶制御装置の第2の記憶領域に記憶されている。第1のリクエストは、その実体が第2の記憶領域に存在する第1の記憶領域を対象とするため、要求された所定のデータ処理を第2の記憶制御装置に実行させることができる。上位装置からのリクエストが、例えば、第1の記憶制御装置のステータス要求や制御情報のバックアップ処理等のように、第1の記憶制御装置自身で処理すべき(あるいは、処理する方が好ましい)データ処理である場合は、第2の記憶制御装置に任せずに、第1の記憶制御装置自らが処理して上位装置に応答することができる。

【0014】

本発明の一態様では、第1の記憶制御装置は、第1の記憶領域と第2の記憶領域との対応関係を示す記憶領域対応情報を保持し、該記憶領域対応情報に基づいて上位装置に第1の記憶領域を仮想的に提供するものであり、第1の制御手段は、記憶領域対応情報に基づいて、第1のリクエストによる第1の記憶領域を対象とするデータ処理を実行可能となっている。

記憶領域対応情報は、第1の記憶領域として割り当てられた第2の記憶領域の対応関係を示す情報であり、例えば、マッピングテーブル等として第1の記憶制御装置内の記憶手段(半導体メモリ等)に保持されている。記憶領域対応情報は、例えば、記憶システムの構成を定義等する際に、オペレータの手動操作により、または自動的な処理により、生成して記憶することが可能である。

【0015】

本発明の一態様では、第2のリクエストは、第1のリクエストと同様のデータ構造を有するように構成されている。

第1のリクエストと第2のリクエストとを同様のデータ構造とすることにより、第2のリクエストを受信した第2の記憶制御装置は、上位装置から直接命令されたと同様にデータ処理を行うことができる。つまり、第2のリクエストを第1のリクエストと異なるデータ構造にすると、第2の記憶制御装置に第2のリクエストを受信し解釈するための機能を追加する必要があるが、第1のリクエストと第2のリクエストとを同様のデータ構造とすることで、第2の記憶制御装置に特別な機能を追加等することなく、第2の記憶制御装置を有效地に利用することができる。

【0016】

本発明の一態様では、第1の制御手段は、第2のリクエストを第2の記憶制御装置に送信する前に、第2のリクエストに係る所定のデータ処理を第2の記憶制御装置が実行可能か否かを確認する。

第2のリクエストを送信する前に、第2のリクエストが要求する所定のデータ処理を第2の記憶制御装置が実行可能か確認することにより、無駄なリクエストを第2の記憶制御装置に送信するのを未然に防止することができる。また、第2の記憶制御装置が所定の処理を行えず、第1の記憶制御装置が実行する場合、無駄なリクエストを送信して第2の記憶制御装置からエラー応答が帰ってくるのを待たずに、直ちに所定の処理を実行することができる。

【0017】

本発明の一態様では、第1の記憶制御装置は、第2の記憶制御装置で実行可能なデータ処理機能を示す機能管理情報を保持し、第1の制御手段は、機能管理情報に基づいて、第2のリクエストに係る所定のデータ処理を第2の記憶制御装置で実行可能か否かを判断するようになっている。

機能管理情報は、例えば、データのバックアップ、ペアボリューム間のコピー、ミラーリング等の各種機能毎に、その機能を実行可能か否か（利用可能か否か）を示す情報を対応付けることにより構成することができる。あるいは、第2の記憶制御装置で実行可能な機能のみを機能管理情報で管理してもよい。

そして、機能管理情報は、記憶システムの構成を定義する際に、手動で又は自動的に生成することができる。

【0018】

本発明の一態様では、第1の記憶制御装置及び第2の記憶制御装置の両方に通信可能に接続されたバックアップ装置を備え、第1の制御手段は、第1のリクエストに係るデータ処理が第1の記憶領域に記憶されている情報をバックアップ装置に転送して記憶させるバックアップ処理である場合、第2の記憶制御装置がバックアップ処理を実行可能か否かを判断し、第2の記憶制御装置がバックアップ処理を実行可能と判断した場合は、第1のリクエストに含まれる第1の記憶領域におけるアドレスを第2の記憶領域におけるアドレスに変換することにより、第2のリクエストを生成して第2の記憶制御装置に送信するものであり、第2の制御手段は、第2のリクエストに基づいて、第2の記憶領域に記憶されている情報をバックアップ装置に転送して記憶させる。

【0019】

記憶システムが、データのバックアップを取るためのバックアップ装置を備えている場合、上位装置は、定期的にまたは不定期に、バックアップ処理を要求する第1のリクエストを第1の記憶制御装置に送信する。第1のリクエストを受信した第1の記憶制御装置（第1の制御手段）は、第2の記憶制御装置がバックアップ処理を実行可能か否かを判断し、実行可能であると判断した場合は、第2のリクエストを生成して第2の記憶制御装置に送信する。第2のリクエストを受信した第2の記憶制御装置は、上位装置から直接命令を受けた如く、第2の記憶領域に記憶されている情報（データや制御情報）を、バックアップ装置に転送して記憶させる。従って、第1の記憶制御装置は、自身が受けたバックアップ要求を第2の記憶制御装置に肩代わりさせることができ、その分、自身の情報処理資源を上位装置への他のサービス提供に使用することができる。

【0020】

本発明の一態様では、第1の記憶制御装置は、第1の記憶領域とペアを構成する第1のセカンダリ記憶領域を仮想的にさらに提供するものであり、第2の記憶制御装置は、第2の記憶領域とペアを構成する第2のセカンダリ記憶領域をさらに有し、第1の制御手段は、第1のリクエストに係る所定のデータ処理が第1の記憶領域に記憶されている情報を第1のセカンダリ記憶領域にコピーする内部コピー処理である場合、第2の記憶制御装置が内部コピー処理を実行可能か否かを判断し、第2の記憶制御装置が内部コピー処理を実行可能と判断した場合は、第1のリクエストに含まれる第1の記憶領域におけるアドレスを第2の記憶領域におけるアドレスに変換することにより、第2のリクエストを生成して第2の記憶制御装置に送信するものであり、第2の制御手段は、第2のリクエストに基づいて、第2の記憶領域に記憶されている情報を第2のセカンダリ記憶領域にコピーさせるようになっている。

【0021】

第1の記憶制御装置及び第2の記憶制御装置は、それぞれ記憶領域を2つずつ備えている。一方の記憶領域はプライマリ記憶領域であり、他方の記憶領域はセカンダリ記憶領域である。プライマリ記憶領域とセカンダリ記憶領域とはペアをなし、それぞれ同一のデータを記憶するものである。上位装置が第1の記憶制御装置に対して、プライマリの第1の記憶領域に記憶されている情報をセカンダリの第1の記憶領域にコピーさせる内部コピー処理を要求した場合、第1の制御手段は、内部コピー処理を第2の記憶制御装置が実行可能か否かを判断し、実行可能であると判断した場合は、第1のリクエスト中のアドレスを第2の記憶制御装置用のアドレスに変換して、第2のリクエストを生成する。即ち、第1のリクエストは、第1の記憶制御装置の第1の記憶領域を対象とするため、第1のリクエスト中に含まれるアドレスは、第1の記憶領域における特定の記憶空間を示すアドレスと

なっている。そこで、第1の制御手段は、第1の記憶領域におけるアドレスを第2の記憶領域における対応するアドレスに変換することにより、第2のリクエストを生成する。これにより、第2の記憶制御装置は、あたかも自身が直接上位装置から命令を受けた場合のように、実記憶領域であるプライマリの第2の記憶領域に記憶されている情報を、実記憶領域であるセカンダリの第2の記憶領域に記憶させる。従って、第1の記憶制御装置は、内部コピー処理を第2の記憶制御装置に代行させて、自身の情報処理資源を他のサービスに使用することができる。

【0022】

本発明の一態様では、第1の記憶制御装置及び第2の記憶制御装置が設けられる正サイトと対をなし、正サイトと通信可能に接続された副サイトを有し、該副サイトは、他の第1の記憶制御装置及び他の第2の記憶制御装置を備え、他の第1の記憶制御装置は、他の第2の記憶制御装置が管理する他の第2の記憶領域を、自己の管理する他の第1の記憶領域として仮想的に提供するものであり、正サイトの第1の制御手段は、第1のリクエストに係る所定のデータ処理が第1の記憶領域に記憶されている情報を副サイトの他の第1の記憶領域にコピーする外部コピー処理である場合、第2の記憶制御装置及び他の第2の記憶制御装置の両方が外部コピー処理を実行可能か否かを判断し、各第2の記憶制御装置がそれぞれ外部コピー処理を実行可能と判断した場合は、第1のリクエストに対応する第2のリクエストを生成して第2の記憶制御装置に送信させ、第2の制御手段は、第2のリクエストに基づいて、第2の記憶領域に記憶されている情報を他の第2の記憶領域にコピーさせることにより外部コピー処理を実行する。

【0023】

記憶システムは、正サイト（プライマリサイト、ローカルサイト）と、正サイトとは離れた場所に設置される副サイト（セカンダリサイト、リモートサイト）との2つのサイトから構成することができる。正サイトは、上位装置と第1の記憶制御装置と第2の記憶制御装置とを備え、副サイトは、他の第1の記憶制御装置及び他の第2の記憶制御装置を備える。副サイトは、正サイトのバックアップ用サイトであり、正サイトの第1の記憶制御装置と副サイトの他の第1の記憶制御装置とはペアを構成する。そして、正副の各サイトにおいて、第1の記憶制御装置（他の第1の記憶制御装置）は、第2の記憶制御装置（他の第2の記憶制御装置）が現実に提供する実際の記憶領域（第2の記憶領域、他の第2の記憶領域）を、自身の記憶領域（第1の記憶領域、他の第1の記憶領域）であるかのように見せかけている。従って、正副の各サイトにおける第2の記憶制御装置と他の第2の記憶制御装置も互いにペアを構成することになる。

【0024】

正サイトに設置された第1の記憶制御装置が上位装置に仮想的に提供する第1の記憶領域は、プライマリボリュームであり、副サイトに設置された第1の記憶制御装置が仮想的に提供する第1の記憶領域は、プライマリボリュームとペアをなすセカンダリボリュームである。上位装置は、定期的にまたは不定期に、正サイトの第1の記憶制御装置に向けて外部コピー処理を要求する。外部コピー処理とは、正サイトの情報を副サイトにコピーさせる処理である。第1のリクエストにより外部コピー処理が要求されると、正サイトに属する第1の制御手段は、正副両サイトの第2の記憶制御装置が外部コピー処理を実行可能であるか否かを判断する。両サイトにそれぞれ設置された第2の記憶制御装置が外部コピー処理を実行可能である場合、第1のリクエストに対応する第2のリクエストを生成し、この第2のリクエストを正サイトに設置された第2の記憶制御装置に送信させる。これにより、正サイトに設置された第2の記憶制御装置は、第2の記憶領域に記憶されているデータを読み出して、副サイトに設置された第2の記憶制御装置に送信し、副サイトの第2の記憶領域に記憶させる。ここで、第1のリクエスト中で明示されているコピー元アドレスやコピー先アドレスを、正副両サイトの第2の記憶領域におけるアドレスに変換することにより、第2のリクエストが生成される。これにより、実際にデータや制御情報を記憶している正副両サイトの第2の記憶制御装置間で外部コピー処理が行われ、正副両サイトの第1の記憶制御装置は、外部データ処理に直接関与しない。このため、正サイトに設置

された第1の記憶制御装置が、外部コピー処理を実行するために情報処理資源を消費することなく、その分の情報処理資源を用いて上位装置にサービスを提供できる。

【0025】

ここで、副サイトに置かれる他の第1の記憶制御装置は、他の第2の記憶制御装置で実行可能なデータ処理機能を示す他の機能管理情報を保持しており、正サイトの第1の制御手段は、第2のリクエストを第2の記憶制御装置に送信する前に、副サイトの他の第1の記憶制御装置に問い合わせることにより、他の第2の記憶制御装置が外部コピー処理を実行可能か否かを判断することができる。

【0026】

つまり、副サイトに設置された他の第2の記憶制御装置が実行可能なデータ処理機能の情報は、副サイト用の機能管理情報として、副サイトに設置された他の第1の記憶制御装置が保持している。そこで、正サイトに設置された第1の記憶制御装置は、第2のリクエストを送信する前に、副サイトに設置された他の第1の記憶制御装置に問合せて、副サイトの他の第2の記憶制御装置が外部コピー処理を実行可能であるか否かを判断する。なお、正サイトに設置された第2の記憶制御装置で実行可能なデータ処理機能の情報は、正サイト用の機能管理情報として、正サイトに設置された第1の記憶制御装置が保持する。

【0027】

本発明の一態様では、第1の記憶制御装置は、第1の記憶領域と第2の記憶領域との対応関係を示す記憶領域対応情報を保持し、他の第1の記憶制御装置は、他の第1の記憶領域と他の第2の記憶領域との対応関係を示す他の記憶領域対応情報を保持し、第1の制御手段は、第2のリクエストを第2の記憶制御装置に送信する場合に、各記憶領域対応情報を送信する。

これにより、正サイトに設置された第2の記憶制御装置は、コピー元及びコピー先の記憶空間に関する情報を把握して、外部コピー処理を行うことができる。

【0028】

本発明の一態様では、第1の記憶制御装置は、外部コピー処理中に上位装置により更新された第1の記憶領域の情報に関する情報を保持する更新位置情報保持手段を備え、第1の制御手段は、外部コピー処理が完了した場合に、更新位置情報保持手段に基づいて、第1の記憶領域において更新された情報を他の第1の記憶領域に記憶させるべく、第2のリクエストを生成して、第2の記憶制御装置から更新された情報を読み出して、読み出した情報を送信する。

即ち、正副両サイトにそれぞれ設置された第2の記憶制御装置間で外部コピー処理が行われている間に、上位装置が、正サイトに設置された第1の記憶制御装置にアクセスしてデータを更新させる場合がある。そこで、外部コピー処理中に更新された情報に関する情報（例えば、更新された論理ロックアドレスの情報等）を更新位置情報保持手段に記憶させておき、外部コピー処理が完了した後で、更新された情報を副サイトの第2の記憶領域に反映させる。

【0029】

本発明の別の観点に従う制御方法は、第1の記憶制御装置と第2の記憶制御装置とを通信可能に接続して構成され、上位装置からのリクエストに応じたデータ処理を行う記憶システムを制御するための制御方法であって、第1の記憶制御装置は、上位装置から第1のリクエストを受信するステップと、受信した第1のリクエストに係る所定のデータ処理を第2の記憶制御装置が実行可能か否かを判断するステップと、第2の記憶制御装置が実行可能と判断した場合は、第1のリクエストに対応する第2のリクエストを生成するステップと、生成された第2のリクエストを第2の記憶制御装置に送信させるステップとを実行し、第2の記憶制御装置は、第1の記憶制御装置から第2のリクエストを受信するステップと、受信した第2のリクエストに基づいて、所定のデータ処理を行うステップとを実行する。

【0030】

本発明の更に別の観点に従う記憶制御装置は、第2の記憶制御装置及び上位装置と通信

可能に接続され、上位装置からのリクエストに応じたデータ処理を行う記憶制御装置であって、上位装置から第1のリクエストを受信する受信手段と、受信した第1のリクエストに係わる所定のデータ処理を第2の記憶制御装置が実行可能か否かを判断する判断手段と、第2の記憶制御装置が実行可能と判断した場合は、第1のリクエストに対応する第2のリクエストを生成するリクエスト手段と、生成された第2のリクエストを第2の記憶制御装置に送信する送信手段と、を備えて構成される。

【0031】

本発明に従うプログラムは、第2の記憶制御装置及び上位装置と通信可能に接続され、上位装置からのリクエストに応じたデータ処理を行う第1の記憶制御装置を制御するプログラムであって、上位装置から受信した第1のリクエストに係る所定のデータ処理を第2の記憶制御装置が実行可能か否かを判断する機能と、第2の記憶制御装置が実行可能と判断した場合は、第1のリクエストに対応する第2のリクエストを生成する機能と、生成された第2のリクエストを第2の記憶制御装置に送信させる機能とを、第1の記憶制御装置のコンピュータ上に実現させる。

【発明を実施するための最良の形態】

【0032】

以下、図1～図22に基づき、本発明の実施形態を説明する。

【0033】

本発明では、以下に詳述するように、第1の記憶制御装置は、第2の記憶制御装置が提供する実記憶領域を、自身の記憶領域のように見せかけてホストコンピュータに仮想的に提供すると共に、第2の記憶制御装置で実行可能な処理については第2の記憶制御装置に任せられるようになっている。即ち、本発明では、分散された物理的な記憶資源を第1の記憶制御装置に論理的に集中して管理可能とし、さらに、集中管理可能な状態下で、負荷の大きな特定の処理を分散して実行させる。

【0034】

ここで、第1の記憶制御装置は、複数種類の動作モードを備えることができる。第1のモードは、本発明に特有なモードであり、ホストコンピュータに対して仮想的な記憶領域を提供すると共に、第2の記憶制御装置で実行可能な処理は第2の記憶制御装置に委ねるモードである。第2のモードは、本発明の前提的な構成におけるモードであり、ホストコンピュータに対して仮想的な記憶領域を提供すると共に、第2の記憶制御装置で実行可能な処理を第1の記憶制御装置で処理するモードである。第3のモードは、従来から知られたモードであり、仮想的な記憶領域を提供せず、実際の記憶領域について第1の記憶制御装置が処理を行うモードである。本発明に従う記憶システムは、少なくとも第1のモードを備える。さらに、第2のモード、第3のモードを備えることもできる。複数種類の動作モードを備える記憶システムは、従来より知られているものではなく、本発明の特徴の一つとなる。

【実施例1】

【0035】

【全体構成の概要】

まず、図1～図9に基づいて、本発明に従う実施例の構成を説明する。図1は、本実施例による記憶システムの要部の構成を示すブロック図である。

【0036】

ホスト装置1は、例えば、C P U (Central Processing Unit) やメモリ等の情報処理資源を備えたコンピュータ装置であり、例えば、パーソナルコンピュータ、ワークステーション、メインフレーム等として構成される。ホスト装置1は、例えば、キーボードスイッチやポインティングデバイス、マイクロフォン等の情報入力装置(図示せず)と、例えば、モニタディスプレいやスピーカー等の情報出力装置(図示せず)とを備えている。ホスト装置1には、R A I Dを管理するためのR A I Dマネージャやデータベースマネジメントプログラム等が実装されている。

【0037】

ホスト装置1は、通信ネットワーク2を介して、後述する第1の記憶制御装置10及び第2の記憶制御装置200にそれぞれ双方向通信可能に接続されている。通信ネットワーク2としては、例えば、LAN、SAN、インターネット、専用回線、公衆回線等を場合に応じて適宜用いることができる。LANを介するデータ通信は、例えば、TCP/IP (Transmission Control Protocol/Internet Protocol) プロトコルに従って行われる。ホスト装置1がLANを介して第1の記憶制御装置10等に接続される場合、ホスト装置1は、ファイル名を指定してファイル単位でのデータ入出力を要求する。ホスト装置1がSANを介して第1の記憶制御装置10等に接続される場合、ホスト装置1は、ファイバチャネルプロトコルに従って、複数のディスク記憶装置（ディスクドライブ）により提供される記憶領域のデータ管理単位であるブロックを単位としてデータ入出力を要求する。

【0038】

なお、ホスト装置1と第1の記憶制御装置10との間、第1の記憶制御装置10と第2の記憶制御装置200との間、ホスト装置1と第2の記憶制御装置200との間は、それぞれ別々の通信ネットワークで接続してもよいし、図示のように共通の通信ネットワーク2で接続してもよい。第1の記憶制御装置10とホスト装置1との間の通信ネットワークが、第1の記憶制御装置10と第2の記憶制御装置200との間の通信ネットワークと相違する場合は、プロトコル変換等を行って第1の記憶制御装置10から第2の記憶制御装置200に第2のリクエストを送信すればよい。

【0039】

第1の記憶制御装置10は、後述のように、例えば、ディスクアレイ装置やファイバチャネルスイッチとして構成されるコンピュータシステムである。第1の記憶制御装置10は、第2の記憶制御装置200が提供する物理的な記憶領域上に設定される論理的な記憶領域（論理ボリューム（Logical Unit））を、第1の記憶制御装置自身が提供するかのように見せかけて、ホスト装置1に仮想的に提供する。

以下の説明では、論理ボリュームをLU、第1の記憶制御装置10が仮想的に提供するLUを仮想的LUと呼ぶ場合がある。また、現実に存在するLUを実LUと称する場合がある。図中では、仮想的LUを仮想線（二点鎖線）で示し、実LUを実線で示す。なお、第1の記憶制御装置10をディスクアレイ装置から構成する場合、第1の記憶制御装置10は仮想的LUと実LUとの両方を有することになる。第1の記憶制御装置10内の実LUは、その物理的な実記憶領域が第1の記憶制御装置10内に、あるいは第1の記憶制御装置10の直接的な管理下にある仮想的なLUであると捉えることもできる。このように考えるならば、記憶領域の実体が第2の記憶制御装置200側にある仮想的LUを仮想的外部LUと、逆に記憶領域の実体が第1の記憶制御装置10側にある仮想的LUを仮想的内部LUと称することも可能である。

【0040】

このように、第1の記憶制御装置10は、第2の記憶制御装置200のLUを仮想的LUとしてホスト装置1に提供するものであるため、第1の記憶制御装置10自身は、実LUを備えている必要はない。従って、物理的な記憶領域を提供するディスクドライブを備えないファイバチャネルスイッチであっても、必要な情報処理能力を有する限り、第1の記憶制御装置10として用いることができる。

【0041】

第1の記憶制御装置10は、ホスト装置1や第2の記憶制御装置200との間で通信を行うためのポート11、12と、コマンド制御部210を備えている。また、第1の記憶制御装置10をディスクアレイ装置から構成する場合、第1の記憶制御装置10は記憶装置70を備える。記憶装置70は、例えば、ハードディスク、フレキシブルディスク、磁気テープ、半導体メモリ、光ディスク等のような装置を複数備えて構成され、これら物理的な記憶領域上にLU71が設定される。また、第1の記憶制御装置10は、第2の記憶制御装置200の実LUをマッピングしてなる仮想的LU72を提供する。図1中では、説明の便宜上、記憶装置70内に実LU71及び仮想的LU72を含めている。記憶装置70は、第1の記憶制御装置10と直接的に接続することもできるし、あるいは、通信ネ

ットワークを介して接続してもよい。また、記憶装置200と第1の記憶制御装置100とを一体化するすることもできる。

【0042】

各ポート11, 12は、データの送受信を行うものである。一方のポート11は、ホスト装置1からのリクエストを受信するターゲットポート(TP)であり、他方のポート12は、第2の記憶制御装置200にリクエストを送信するイニシエータポート(IP)である。両ポート11, 12の構造は同一であり、データ通信上の役割に応じて、ターゲットポートまたはイニシエータポートとなる。

【0043】

コマンド制御部100は、第1の記憶制御装置100が有する情報処理資源(CPU、メモリ、入出力回路等)やソフトウェアにより実現される。コマンド制御部100は、例えば、第1の記憶制御装置100のメインコントローラ(図示せず)により実現することもできるし、チャネルアダプタやディスクアダプタの連携により実現することもできる。コマンド制御部100は、コマンド解析部110と、判断部120と、コマンド生成部130と、コマンドテーブルTC、マッピングテーブルTM及び機能管理テーブルTFを備えている。

【0044】

コマンド解析部110は、ホスト装置1からポート11を介して受信したコマンド(リクエスト)をコマンドテーブルTCに基づいて解析するものである。コマンドテーブルTCには、予め各種のコマンドコードが登録されており、コマンドテーブルTCを参照することにより、ホスト装置1からのリクエスト中に含まれるコマンドコードが何を要求する命令であるかを判別できる。

【0045】

判断部120は、ホスト装置1から受信したリクエストに係わるデータ処理を、第2の記憶制御装置200で実行可能か否かを判断するものである。判断部120は、機能管理テーブルTFを参照することにより、要求されたデータ処理機能を第2の記憶制御装置200が備えているか否かを判定する。機能管理テーブルTFの詳細は後述するが、機能管理テーブルTFは、例えば、記憶システムの構成を定義等する際に生成されるもので、第2の記憶制御装置200に実装されているサポート機能が登録されている。

【0046】

コマンド生成部130は、ホスト装置1から受信したリクエストを第2の記憶制御装置200用に変換するものである。コマンド生成部130は、コマンドテーブルTC及びマッピングテーブルTMを参照することにより、ホスト装置1が要求するデータ処理を第2の記憶制御装置200が実行できるようにリクエストを生成する。マッピングテーブルTMは、第2の記憶制御装置200のLUと第1の記憶制御装置100の仮想的LUとの対応関係を示すものである。コマンド生成部130は、マッピングテーブルTMを参照することにより、仮想的LUの記憶空間を対象とするアドレスを実LUの記憶空間を対象とするアドレスに変換する。

【0047】

このように、コマンド制御部100は、ホスト装置1から受信した第1のリクエストを解析し、第1のリクエストで要求されたデータ処理(サポート機能)を第2の記憶制御装置200が実行可能であるか否かを判断する。そして、第2の記憶制御装置200で実行可能と判断した場合、コマンド制御部100は、第1のリクエストに対応する第2のリクエストを生成し、この第2のリクエストをポート12から通信ネットワーク2を介して第2の記憶制御装置200に送信する。ホスト装置1から要求されたデータ処理を第2の記憶制御装置200が実行できない場合は、要求されたデータ処理を第1の記憶制御装置100が実行する。

【0048】

[第1の記憶制御装置をディスクアレイ装置に適用した場合]

図2は、第1の記憶制御装置100をディスクアレイ装置として構成した場合の具体例を

示すブロック図である。第1の記憶制御装置10は、それぞれ後述するように、例えば、複数のチャネルアダプタ20と、キャッシングメモリ40と、共有メモリ50と、複数のディスクアダプタ60と、記憶装置70等とから構成することができる。また、図示を省略するが、第1の記憶制御装置10は、例えば、第1の記憶制御装置10の全体動作を制御するためのMPU(Micro Processing Unit)と、環境設定を行ったり各種ステータスを管理等するための保守管理用端末とを備えることができる。

【0049】

チャネルアダプタ(CHA)20は、ホスト装置1との間のデータ通信を行うものである。各チャネルアダプタ20は、ホスト装置1等と通信を行うための通信ポート21と、受信データを転送するための転送部22と、チャネルアダプタ内の動作を制御するマイクロプロセッサ(MPと略記)23と、メモリ24とを備えている。マイクロプロセッサ23等によって、ホスト装置1から受信した各種コマンドを解釈して処理するためのコマンド制御部100を実現している。あるいは、装置全体を統括するMPUによってコマンド制御部100を実現してもよく、さらには、MPUとチャネルアダプタ20との連携、MPUとチャネルアダプタ20及びディスクアダプタ60との連携、チャネルアダプタ20とディスクアダプタ60との連携によってコマンド制御部100を実現してもよい。

【0050】

各チャネルアダプタ20には、それぞれを識別するためのネットワークアドレス(例えば、IPアドレスやWWN(World wide Name))が割り当てられており、各チャネルアダプタ20は、それが個別にDAS(Direct Attached Storage)として、あるいはNAS(Network Attached Storage)として振る舞うようになっている。つまり、各チャネルアダプタ20は、各ホスト装置1からの要求をそれぞれ個別に受け付けることができる。各チャネルアダプタ20は、ポート21と、転送部22と、マイクロプロセッサ23及びメモリ24からなる制御回路を複数備えて構成することができる。

【0051】

各マイクロプロセッサ23は、ホスト側共有メモリアクセス回路31を介して共有メモリ50にそれ接続されており、共有メモリ50に制御情報を書き込んだり、共有メモリ50に書き込まれた制御情報を参照する。各転送部22は、ホスト側キャッシングメモリアクセス回路32を介してキャッシングメモリ40にそれ接続されており、ホスト装置1等から受信したデータをキャッシングメモリ40に記憶させ、また、キャッシングメモリ40に記憶されたデータを読み出してホスト装置1等に送信するようになっている。キャッシングメモリ40及び共有メモリ50は、各チャネルアダプタ20及び各ディスクアダプタ60によって共有される記憶手段である。キャッシングメモリ40には主としてデータが記憶され、共有メモリ50には主として制御情報やコマンド等が記憶される。また、共有メモリ50には、ワーク領域も設定される。上述したコマンドテーブルTC、マッピングテーブルTM、機能管理テーブルTFは、例えば、共有メモリ50に記憶させることもできるし、あるいは、キャッシングメモリ40に記憶させることもできる。

【0052】

各ディスクアダプタ(DKA)60は、記憶装置70のディスク装置群73との間のデータ入出力を管理するものである。ディスクアダプタ60は、チャネルアダプタ20がホスト装置1から受信したデータを、ホスト装置1からの書き込み要求に従って、ディスク装置群73の所定のアドレスに書き込む。このとき、各ディスクアダプタ60は、論理ボリュームにおける論理アドレスを、物理的ディスクにおける物理アドレスに変換する。ディスクアダプタ60は、ディスク装置群73がRAIDにより管理されている場合は、RAID構成に従ったデータアクセスを行い、ディスク装置群73に記憶されたデータの複製管理制御やバックアップ制御も行うことができる。さらに、ディスクアダプタ60は、災害発生時のデータ消失防止等を目的として、プライマリサイトのデータの複製をセカンダリサイトに記憶させる制御(レプリケーション機能またはリモートコピー機能)等も行うことができる。

【0053】

各ディスクアダプタ60は、通信ポート61と、転送部62と、マイクロプロセッサ63と、メモリ64とからなる制御回路をそれぞれ複数組備えることができる。各通信ポート61は、記憶装置70のディスク装置群73との間でデータ通信を行う。各転送部62は、デバイス側キャッシュメモリアクセス回路34を介してキャッシュメモリ40に接続されており、キャッシュメモリ40に書き込まれたデータをディスク装置群73に転送したり、ディスク装置群73から読み出したデータをキャッシュメモリ40に書き込む。マイクロプロセッサ63は、デバイス側共有メモリアクセス回路33を介して共有メモリ50に接続されており、共有メモリ50に書き込まれた制御情報やコマンドを参照等できるようになっている。

【0054】

ホスト側共有メモリアクセス回路31及びデバイス側共有メモリアクセス回路33と、ホスト側キャッシュメモリアクセス回路32及びデバイス側キャッシュメモリアクセス回路34とは、例えば、高速スイッチング動作によってデータ伝送を行う超高速クロスバスイッチ等のような高速バスとして構成することができる。

【0055】

図3は、第1の記憶制御装置10をホスト装置1から見た場合の論理的な構成の要部を示すブロック図である。第1の記憶制御装置10は、ホスト装置1に対して、2種類のLUを提供する。一つは、第1の記憶制御装置10が直接的に管理する記憶装置70の物理的なディスク装置群73により提供される実LU71である。他の一つは、第2の記憶制御装置200が管理する記憶装置220のディスク装置群221により提供される実LUを、第1の記憶制御装置10のLUとして提供する仮想的LU72である。図3に示すように、実LU71及び仮想的LU72のいずれも複数設けることができる。各LU71, 72は、それぞれ複数のディスク装置群から構成される。

【0056】

[第1の記憶制御装置をファイバチャネルスイッチに適用した場合]

次に、図4は、第1の記憶制御装置10をインテリジェント化されたファイバチャネルスイッチとして構成した場合の具体例を示すブロック図である。ディスクアレイ装置として構成する場合と区別するために、図中では、符号10に(SW)を附加してある。第1の記憶制御装置10は、複数のチャネルアダプタ20と、共有メモリアクセス回路31, 33と、キャッシュメモリアクセス回路32, 34と、キャッシュメモリ40と、共有メモリ50と、コントローラ80とを備えて構成される。チャネルアダプタ20等の詳細は、図2と共に述べたので割愛する。コントローラ80は、全体の動作を統括して制御するものであり、MPUやメモリ等を備えている。図2に示すものとの大きな相違点は、ファイバチャネルスイッチとして構成される第1の記憶制御装置10は、直接的に管理する記憶装置70を備えていない点である。

【0057】

図5は、ファイバチャネルスイッチとして構成される第1の記憶制御装置10をホスト装置1から見たときの論理的な構成の要部を示すブロック図である。図に示すように、第1の記憶制御装置10は、実LUを有していないが、少なくとも一つ以上の仮想的なLU72を備えている。上述のように、仮想的LU72の実体は、第2の記憶制御装置200の記憶装置220に存在する。

【0058】

[仮想的LUへのデータアクセス]

次に、図6を参照して、仮想的LU72へのデータアクセスについて説明する。図6は、第1の記憶制御装置10をファイバチャネルスイッチとして構成した場合を示す。

【0059】

ホスト装置1が仮想的LU72に対してデータの書込みまたはデータの読み出しを要求する場合、ホスト装置1は第1の記憶制御装置10に向けてコマンドCHを発行する。このコマンドCHには、第1の記憶制御装置10を特定するための情報(ポートID、WWN等)と、命令の種類を示すコマンドコード(ライト命令、リード命令等)と、対象とす

るデータの読み出しアドレス（リード命令の場合）等が含まれている。ホスト装置1からのコマンドCHは、SAN等の通信ネットワーク2を介して、第1の記憶制御装置10のターゲットポート11に受信され、コマンド制御部100に入力される。

【0060】

コマンド制御部100は、受領したコマンドCHを解析し、マッピングテーブルTMを参照することにより、ホスト装置1から要求されたデータ処理を行う。第1の記憶制御装置10をファイバチャネルスイッチから構成する場合、マッピングテーブルTMは、例えば、図9(a)に示すように構成できる。マッピングテーブルTMは、例えば、仮想的LU(仮想的な論理ボリューム)72を識別するためのボリュームID(Vol ID)と、仮想的LU72に設定された論理ブロックアドレス(BLK ADDR)と、仮想的LU72に対応付けられる実LU222を有する第2の記憶制御装置200を識別するための装置IDと、実LU222に対応するポートを識別するためのポートIDと、実LU222を識別するためのボリュームIDと、仮想的LU72の論理ブロックアドレスに対応して実LU222に設定されている論理ブロックアドレスとを、それぞれ対応付けることにより構成することができる。従って、マッピングテーブルTMを参照することにより、仮想的LU72の特定の論理ブロックアドレスが、いずれの第2の記憶制御装置200が提供する実LU222のどの論理ブロックアドレス(以下、「アドレス」と略記)に対応するかを把握できる。このマッピングテーブルTMは、例えば、記憶システムが構成され、論理ボリュームが登録されるときに、手動または自動的に登録される。

ここで、仮想LUと実LUとの間のマッピングとしては、全てのアドレスについてマッピングテーブルを使用する場合のほか、テーブルの参照に加えて幾らかの計算により対応するアドレスを求める場合とがある。後者の方法では、前者に比べてマッピングテーブルのサイズを小さくすることができる。本発明におけるマッピングでは、テーブル参照及び計算によって、対応するアドレスを求めることができるようになっている。

【0061】

さて、ホスト装置1が、仮想的LU72のアドレスBbからBcの範囲にデータを書き込ませるためのコマンドCHを発行したとする。コマンド制御部100は、コマンドCHを受信すると、指定されたアドレスに基づいてマッピングテーブルTMを参照する。これにより、コマンド制御部100は、仮想的LU72とポート201(ポートID=TP2)からアクセス可能なボリュームID α 2で特定される実LU222とが対応付けられていることを知ると共に、仮想的LU72におけるアドレスBb-Bcが実LU222のアドレスBb2-Bc2に対応していることを把握する。そこで、コマンド制御部100は、ホスト装置1から受信したコマンドCHの内容を一部書き換えることにより、コマンドCSを生成する。即ち、コマンド制御部100は、ホスト装置1から受信したコマンドCHに含まれるボリュームIDとアドレスとを、マッピングテーブルTMに基づいて、実LU222のボリュームIDとアドレスに書き換えることにより、コマンドCSを生成する(Vol ID α \rightarrow Vol ID α 2、BLK ADDR Bb-Bc \rightarrow BLK ADDR Bb2-Bc2)。そして、コマンド制御部100は、生成したコマンドCSをイニシエータポート12から通信ネットワーク2を介して第2の記憶制御装置200に送信する。なお、以下の場合も同様であるが、アドレス変換を行う場合は、マッピングテーブルTMのみを使用する場合に限らず、マッピングテーブルTMの参照に加えて幾らかの演算を行うことにより、対応するアドレスを求めることができる。

【0062】

第1の記憶制御装置10から送信されたコマンドCSは、第2の記憶制御装置200のターゲットポート201により受信され、コマンド制御部210に引き渡される。コマンド制御部210は、コマンドCSの内容を解析し、物理的に存在する実LU222の指定されたアドレス(Bb2-Bc2)にデータを書き込む。データの書き込みが完了すると、第2の記憶制御装置200は、第1の記憶制御装置10に書き込み完了を報告する。なお、ホスト装置1へ書き込み完了報告を行うタイミングとしては、例えば、第1の記憶制御装置10がコマンドCHを受信した時点(同期式)、または第2の記憶制御装置200から

書き込み完了報告を受けた時点（同期式）のいずれかで行うことができる。同期式の場合は、第2の記憶制御装置200からの応答を待つ時間の分だけ遅延が生じるため、第1の記憶制御装置10と第2の記憶制御装置200とがあまり離れて設置されていない場合に適している。第1の記憶制御装置10と第2の記憶制御装置200とが遠距離で離れている場合は、応答遅延及び伝播遅延の問題から、同期式は一般的に適さず、非同期式が採用される。

【0063】

ホスト装置1が仮想的LU72からデータを読み出す場合も、上述したデータ書き込み時と同様の処理が行われる。ホスト装置1が発行したコマンドCH（リードリクエスト）は、第1の記憶制御装置10のコマンド制御部100によって解析される。コマンド制御部100は、マッピングテーブルTMを参照して、読み出し対象として指定された仮想的LU72のアドレスを実LU222のアドレスに変換してコマンドCSを生成し、コマンドCSを第2の記憶制御装置200に送信する。コマンドCSは、第2の記憶制御装置200のターゲットポート201に受信され、コマンド制御部210により解析される。コマンド制御部210は、指定されたアドレスのデータを実LU222（既にキャッシュメモリに読み出されている場合はキャッシュメモリ）から読み出し、読み出したデータを第1の記憶制御装置10に送信する。第1の記憶制御装置10は、受信したデータをホスト装置1に送信する。ホスト装置1から指定されたアドレスのデータが、第1の記憶制御装置10内のキャッシュメモリ40に既に記憶されている場合は、そのデータをホスト装置1に送信してもよい。

【0064】

次に、図7は、第1の記憶制御装置10をディスクアレイ装置として構成した場合を示す。この場合、第1の記憶制御装置10は、自身が直接的に管理する実LU71と、仮想的LU72とを備えている。仮想的LU72のボリュームIDを α 、実LU71のボリュームIDを β とする。

【0065】

第1の記憶制御装置10が実LU71及び仮想的LU72を備える場合、マッピングテーブルTMは、例えば、図9（b）に示すように構成することができる。図9（b）に示すテーブルのうち、仮想的LU72に関連する部分は図9（a）に示すテーブルと同一である。異なる点は、第1の記憶制御装置10の実LU71に関連する部分である。実LU71は、第1の記憶制御装置10自身が有るので、その装置IDは第1の記憶制御装置10のIDとなる（SD1）。同じ理由から、実LU71にアクセスするためのポートIDには、内部の論理ボリュームであることを示す情報（INTERNAL）が登録される。また、第2の記憶制御装置200のLUとの対応関係は存在しないので、対応するボリュームIDやアドレスについては登録されない。ポートIDフィールド及び内外の区別を示す識別子（INTERNAL）を利用することにより、実LU71と仮想的LU72とを同一構造のマッピングテーブルTMで管理することができる。なお、これに限らず、実LU71と仮想的LU72とを別々のテーブルで管理してもよい。

ホスト装置1から仮想的LU72へのデータの書き込み及び読み出しは、上述したファイバチャネルスイッチの場合と同様であるので、説明を割愛する。

【0066】

次に、図8は、第1の記憶制御装置10をインテリジェント化されたファイバチャネルスイッチとして構成すると共に、第1の記憶制御装置10の仮想的LU72が、第2の記憶制御装置200A, 200Bがそれぞれ提供する実LU222A, 222Bの2つの論理ボリュームから構成される場合を示している。この場合、マッピングテーブルTMは、例えば、図9（c）に示すように構成することができる。一方の第2の記憶制御装置200Aについて、その装置IDをSD2（1）、ポートIDをTP2（1）、ボリュームIDを α 2Aとし、同様に、他方の第2の記憶制御装置200Bについては、その装置IDをSD2（2）、ポートIDをTP2（2）、ボリュームIDを α 2Bとする。

【0067】

図9(c)に示すマッピングテーブルTMに示すように、第1の記憶制御装置10がホスト装置1に対して提供する仮想的LU72(ボリュームID=α)は、ポートID=TP2(1)からアクセス可能なLU222A(ボリュームID=α2A)と、ポートID=TP2(2)からアクセス可能なLU222B(ボリュームID=α2B)とから構成されていることがわかる。このように、第1の記憶制御装置10は、分散された複数のLUをまとめて一つ又は複数の仮想的LUを構築することができる。従って、例えば、使用効率の低い旧型の記憶制御装置を複数台まとめて、一つ又は複数の仮想的LUを構成し、ホスト装置1に提供することにより、記憶システムの記憶資源を再編成して有効に利用することができる。

【0068】

【ダイレクトバックアップ1】

次に、ダイレクトバックアップを行う場合の流れを説明する。ダイレクトバックアップとは、ホスト装置1を介在させずに、記憶制御装置とバックアップ装置との間で直接的にデータのバックアップを取る処理をいう。ダイレクトバックアップは、記憶制御装置が提供するサポート機能の一種である。

【0069】

図10は、一般の方法と考えられるダイレクトアップの仕組みを簡単に示す説明図である。図10に示す記憶制御装置10(N)は、仮想的LUを構築しない普通の記憶制御装置である。バックアップ装置3は、データの全体または一部の複製を記憶する記憶装置である。バックアップ装置3としては、例えば、MO(magneto-optic:光磁気型記憶装置)、CD-R(CD-Recordable:読み書き可能なコンパクトディスク)、DVD-RAM(Digital Versatile Disk-RAM:読み書き可能なDVD)等のディスク系記憶デバイスや、例えば、DAT(Digital Audio Tape)テープ、カセットテープ、オープンテープ、カートリッジテープ等のテープ系記憶デバイス等を用いることができる。図示の例では、テープ系記憶デバイスを想定しているが、これに限定されない。

【0070】

ダイレクトバックアップにおいて、バックアップ装置へのデータの書込みは、2種類に大別することができる。その一つは、データのバックアップであり、他の一つは、制御情報の書込みである。いずれのバックアップを行うかは、ホスト装置1が発行するコマンド(リクエスト)により決定される。データのバックアップを行う場合は、図10中の(a)に示すようなコピーリクエストをホスト装置1から記憶制御装置10(N)に送信する。データのバックアップを行う場合、ホスト装置1から送信されるコマンドCHには、コマンドコードにコピーリクエストを示すリクエストIDが、コピー元アドレスにはバックアップを取るデータの範囲を示すアドレス及びボリュームID(アドレスのみ図示)が、コピー先装置IDにはバックアップ装置3を特定するための装置IDが、それぞれ格納される。一方、制御情報のバックアップを行う場合、コマンドCHには、コマンドコードに制御情報の書込みリクエストを示すリクエストIDが、コピー先装置IDにはバックアップ装置3のIDがそれぞれ格納されると共に、コマンドCH内にはコピー対象である制御情報が含まれる。これらのリクエストの実例としては、例えば、SCSI-3にて規定されるExtended Copyコマンドを挙げることができる。

【0071】

まず先に、データのバックアップを行う場合の動作を説明する。ホスト装置1は、図10(a)に示すようなコマンド(コピーリクエスト)CHを生成し、LU71中のバックアップ対象とする論理ブロックのアドレスを指定する。生成されたコマンドCHは、ホスト装置1から通信ネットワーク2を介して記憶制御装置10(N)に送信される。記憶制御装置10(N)は、ターゲットポート11によってコマンドCHを受信する。受信されたコマンドCHは、コマンド制御部100により解析され、データのバックアップを要求するコピーリクエストであることが認識される。コマンド制御部100は、コマンドCH中のコピー元アドレスを参照することにより、指定された論理ボリュームβのうち指定されたアドレスBa-Bdの論理ブロックのデータを読み出す。コマンド制御部100は、

読み出したバックアップ対象のデータをバックアップ装置3に書き込ませるためのコマンド（書込みリクエスト）を生成し、イニシエータポート12からバックアップ装置3に向けて書込みコマンドを送信する。記憶制御装置10（N）から書込みコマンドを受信したバックアップ装置3は、受信データを所定の位置に書き込み、書き込みが完了した場合は、書き込み完了報告を記憶制御装置10（N）に送信する。

【0072】

次に、制御情報の書き込みを行う場合の動作を説明する。まず、ホスト装置1は、図10（b）に示すような制御情報書き込みコマンドを生成する。ホスト装置1からのコマンドCHを受信した記憶制御装置10（N）は、コマンドCHを解析して内容を把握し、コマンドCHに含まれる制御情報を抽出する。そして、前記同様に、コマンド制御部100は、抽出した制御情報をバックアップ装置3に書き込ませるための書き込みコマンドを生成して、バックアップ装置3に送信する。記憶制御装置10（N）から書き込みコマンドを受信したバックアップ装置3は、制御情報を所定の位置に書き込み、書き込み完了報告を記憶制御装置10（N）に送信する。

【0073】

【ダイレクトバックアップ2】

次に、図11を参照して、ダイレクトバックアップの別の方法を説明する。図11には、第2の記憶制御装置200の実記憶領域をあたかも自己の記憶領域であるかのように見せかけて、ホスト装置1に仮想的に提供する構成において、ダイレクトバックアップを行う場合を示している。

【0074】

図に示すように、例えば、インテリジェント化されたファイバチャネルスイッチから構成される第1の記憶制御装置10と、第2の記憶制御装置200と、ホスト装置1と、バックアップ装置3とは、それぞれ通信ネットワーク2を介して相互通信可能に接続されている。そして、上述の通り、第2の記憶制御装置200が直接管理する実LU222は、マッピングテーブルTMによって、第1の記憶制御装置10の仮想的LU72に割り付けられている。これにより、第1の記憶制御装置10は、第2の記憶制御装置200の実LUを間接的に支配し、仮想的LU72としてホスト装置1に提供している。

【0075】

先にデータのバックアップを行う場合を説明すると、まず、ホスト装置1は、コピーリクエストを示すリクエストIDと、コピー対象のデータを特定するためのボリュームID（ α ）及びアドレス（Ba-Bd）と、コピー先装置を特定するための装置ID（バックアップ装置の装置ID）とを含んだコマンドCHを生成する。次に、ホスト装置1は、通信ネットワーク2を介してコマンドCHを第1の記憶制御装置10に送信すると、このコマンドCHは、第1の記憶制御装置10のターゲットポート11により受信される。コマンド制御部100は、コマンドCHを解析し、コマンドCH中のコピー元アドレスとマッピングテーブルTMとを参照することにより、コピー対象となっている仮想的LU72のアドレスに対応する実LU222のアドレスを検索する。即ち、コマンド制御部100は、ダイレクトバックアップを指定された仮想的LU72に対応する論理ボリュームがターゲットポート201からアクセス可能な実LU222であることを知ると共に、ダイレクトバックアップの範囲が実LU222におけるアドレスBa2-Bd2の範囲であることを知る。

【0076】

次に、コマンド制御部100は、実LU222のアドレスBa2-Bd2に記憶されているデータを読み出すための読み出しコマンド（読み出しリクエスト）を生成し、読み出しコマンドをイニシエータポート12から送信する。読み出しコマンドは、通信ネットワーク2を介して、第2の記憶制御装置200のターゲットポート201に受信される。第2の記憶制御装置200のコマンド制御部210は、読み出しコマンドにより要求された範囲のデータを実LU222から読み出して、第1の記憶制御装置10に送信する。これにより、第1の記憶制御装置10は、ダイレクトバックアップの対象とされたデータを取

得して、データをキャッシュメモリ40に一時格納する。そして、コマンド制御部100は、第2の記憶制御装置200から取得したデータをバックアップ装置3に書き込ませるための書き込みコマンド（コピーリクエスト）を生成し、この書き込みコマンドをコピー先装置として指定されたバックアップ装置3に向けて送信する。バックアップ装置3は、第1の記憶制御装置10から受信した書き込みコマンドに基づいて、受信したデータを所定の位置に格納する。

【0077】

制御情報の書き込みを行う場合を説明する。この場合は、図10と共に説明した一般的な方法と考えられるダイレクトバックアップと同様の動作となる。即ち、ホスト装置1は、制御情報の書き込みリクエストを示すリクエストIDと、コピー先の装置を特定する装置IDと、バックアップ対象の制御情報を含むコマンドCHを生成して、第1の記憶制御装置10に送信する。コマンド制御部100は、受信したコマンドCHを解析し、コマンドCHから抽出された制御情報をバックアップ装置3に書き込ませるための書き込みコマンドを生成して、バックアップ装置3に送信する。バックアップ装置3は、書き込みコマンドを受信すると、制御情報を所定の位置に格納させる。

【0078】

このように、第2の記憶制御装置200の実LU222を第1の記憶制御装置10の記憶領域として仮想的に取り込む構成においては、第1の記憶制御装置10が記憶システムの動作を主導して、ホスト装置1から要求されたデータ処理を実行する（データのダイレクトバックアップの場合）。従って、例えば、第1の記憶制御装置10が間接的に支配下におく第2の記憶制御装置200の数、ホスト装置1から要求される処理内容、通信ネットワーク2の速度等のような諸条件によって相違するが、第1の記憶制御装置10の処理負担は重くなる。そこで、本発明では、以下に述べるように、第1の記憶制御装置10で実行可能な処理の全部または一部を、第2の記憶制御装置200に実行させることにより、負荷分散を図っている。

【0079】

【ダイレクトバックアップ3】

次に、図12～図14に基づいて、さらに別のダイレクトバックアップの方法を説明する。この方法は、ダイレクトバックアップを第2の記憶制御装置200で実行可能な場合に、第2の記憶制御装置200にダイレクトバックアップを行わせる点に特徴がある。

【0080】

データのダイレクトバックアップを希望する場合、ホスト装置1は、ダイレクトバックアップ用のコマンドCHを生成する。このコマンドCHは、上述のように、コピーリクエストを示すリクエストIDと、コピー元アドレスと、コピー先装置の装置IDとが含まれている。ホスト装置1が通信ネットワーク2を介してコマンドCHを送信すると、コマンドCHは第1の記憶制御装置10のターゲットポート11により受信される。

【0081】

コマンド制御部100は、コマンドCHを解析し、リクエストIDに基づいてコマンドテーブルTC（図示省略）を参照することにより、コピーリクエストを受信したこと認識する。コマンド制御部100は、コマンドCHにより指定されたコピー元アドレスに基づいてマッピングテーブルTMを参照し、仮想的LU72のアドレスに対応する実LU22のアドレスを検索する。即ち、コマンド制御部100は、指定された仮想的LU72に対応するボリュームのID（ α 2）を検出すると共に、仮想的LU72の記憶空間におけるアドレス（Ba-Bd）に対応する実LU222の記憶空間におけるアドレス（Ba2-Bd2）を検出する。

【0082】

次に、コマンド制御部100は、機能管理テーブルTFを参照し、第2の記憶制御装置200で実行可能なサポート機能を確認する。図13（a）は、機能管理テーブルTFの一例を示す。図13（a）は、複数の第2の記憶制御装置200がそれぞれ有する機能を管理する場合を示す。機能管理テーブルTFには、例えば、第2の記憶制御装置200を

それぞれ特定するための装置ID（SD2(1)～SD2(n)）と、それぞれの実LU222にアクセスするためのポートID（TP2(1)～TP2(n)）と、各サポート機能（F1～Fn）を実行可能か否かの判別情報とが対応付けられている。実行可能な機能には「可」と記録され、実行不能な機能には「不可」と記録されている。サポート機能としては、例えば、ダイレクトバックアップ、ペアボリュームのコピー、ミラーリング、リモートコピー等を挙げることができる。この機能管理テーブルTFは、例えば、記憶システムが構成されるときに、手動または自動的に登録される。第1の記憶制御装置10に第2の記憶制御装置200が1台だけ接続される場合、図13(a)に示す機能管理テーブルTFのレコードは一つとなる。

【0083】

コマンド制御部100は、機能管理テーブルTFを参照して、第2の記憶制御装置200がダイレクトバックアップ機能をサポートしていることを確認すると、コマンド制御部100は、第2の記憶制御装置200にダイレクトバックアップを担当させるべく、第1のコマンドCHの一部を書き換えることにより、第2のコマンドCSを生成する。具体的には、例えば、図13(b), (c)に示すように、コマンドCH中の「コピー元アドレス」に格納されているボリュームID及びアドレスを、マッピングテーブルTMを参照して、対応する実LU222のボリュームID及び対応するアドレスにそれぞれ書き換えることにより、コマンドCSが生成される。即ち、リクエストの種類を示すリクエストID及びコピー先装置IDは、両コマンドCH, CS間で共通し、ダイレクトバックアップの対象とするデータを特定するための情報のみが、そのデータを実際に記憶している記憶空間のアドレスに変換される。従って、両コマンドCH, CSは、その一部の内容が異なるだけでデータ構造は共通する。

【0084】

このようにしてコマンド制御部100により組み立てられたコマンドCSは、イニシエータポート12から通信ネットワーク2を介して第2の記憶制御装置200のターゲットポート201に到達する。第2の記憶制御装置200のコマンド制御部210は、コマンドCSを解析し、指定された範囲のデータを実LU222から読み出す。そして、コマンド制御部210は、この読み出したデータをバックアップ装置3に書き込ませるための書込みコマンドを生成し、バックアップ装置3に送信する。バックアップ装置3は、第2の記憶制御装置200からの書込みコマンドを受信すると、受信データを所定の位置に記憶させる。

【0085】

図14を参照して第1の記憶制御装置10の動作を説明する。第1の記憶制御装置10がホスト装置1からコマンドCHを受信すると(S1)、コマンド制御部100はコマンドCHを解析し(S2)、機能管理テーブルTFを参照する(S3)。そして、第2の記憶制御装置200でデータのダイレクトバックアップを実行可能と判断した場合は(S4:YES)、第2の記憶制御装置200にデータのダイレクトバックアップを肩代わりさせるためのコマンドCSを生成し(S5)、コマンドCSを第2の記憶制御装置200に送信する(S6)。第2の記憶制御装置200がダイレクトバックアップ機能をサポートしていない場合は(S4:NO)、コマンド制御部100は、読み出しコマンドを生成して第2の記憶制御装置200に送信することにより、第2の記憶制御装置200の実LU222からバックアップ対象のデータを読み出す(S7)。コマンド制御部100は、書込みコマンドを生成してバックアップ装置3に送信することにより(S8)、読み出したデータをバックアップ装置3に記憶させる。

【0086】

なお、制御情報の書き込みを行う場合は、上述した一般的と考えられる方法で書き込まれる。即ち、コマンド制御部100は、受信したコマンドCHから制御情報を抽出し、この抽出した制御情報をバックアップ装置3に書き込むための書込みコマンドを生成し、バックアップ装置3に送信する。

【0087】

以上詳述した通り、本実施例によれば、ホスト装置1から要求されたデータ処理（ダイレクトバックアップ）が第2の記憶制御装置200で実行可能か否かを判定し、第2の記憶制御装置200で実行できる場合には、要求されたデータ処理を第2の記憶制御装置200に実行させることができる。従って、第1の記憶制御装置10に処理が集中するのを防止し、負荷を分散させることができるために、第1の記憶制御装置10の処理負担を軽減することができる。これにより、第1の記憶制御装置10の情報処理資源を他のサービス実現のためにホスト装置1に提供することができ、記憶システム全体の有効利用を図ることができる。

【実施例2】

【0088】

[論理ボリュームコピー1]

次に、図15に基づいて、論理ボリュームのコピーを説明する。本実施例による論理ボリュームコピーは、それぞれ離れて設置された正サイト（プライマリサイト）と副サイト（セカンダリサイト）との間で、論理ボリュームの内容がコピーされる。

【0089】

正サイトには、ホスト装置1と、インテリジェント化されたファイバチャネルスイッチから構成された第1の記憶制御装置10（SW1）と、第2の記憶制御装置200（1）とが含まれており、それぞれ通信ネットワーク2（1）によって相互に通信可能に接続されている。副サイトには、インテリジェント化されたファイバチャネルスイッチから構成された第1の記憶制御装置10（SW2）と、第2の記憶制御装置200（2）とが含まれており、通信ネットワーク2（2）により相互に通信可能に接続されている。正サイトの通信ネットワーク2（1）と副サイトの通信ネットワーク2（2）も相互に接続されている。以下、正副いずれのサイトに設置されたものであるかを明示する場合には、プライマリ側を意味する（1）またはセカンダリ側を意味する（2）の符号を追加し、特に区別する必要が無い場合は省略する。

【0090】

本実施例による論理ボリュームコピーは、コピー元となる正サイトの論理ボリューム（ α ）とコピー先となる副サイトの論理ボリューム（ β ）の組（ペアボリューム）を決めて、正サイトの論理ボリューム（ α ）の記憶内容を副サイトの論理ボリューム（ β ）にコピーすることを意味し、2つの種類に分けて考えることができる。その一つは、コピー元である正サイトの論理ボリューム（ α ）の全データを、コピー先である副サイトの論理ボリューム（ β ）に送信して書き込ませるという、初期コピーである。

他一つは、初期コピー完了後に、ホスト装置1が正サイトの論理ボリューム（ α ）に対して行った更新分のデータのみを、副サイトの論理ボリューム（ β ）に送信して書き込ませるという、更新コピーである。

【0091】

初期コピーは、コピー元となる論理ボリュームが有する全ての論理ブロックの記憶内容をコピー先の論理ボリュームに転送する処理であり、初期コピーを実行する記憶制御装置にとって処理負荷が大きい。そこで、本実施例では、初期コピー時に、第1の記憶制御装置に過大な負荷がかかるのを防止する手段について説明する。

【0092】

まず最初に、図に示す前提的な構成において、正サイトの第1の記憶制御装置10（SW1）が、論理ボリュームのコピーを実行する場合を説明する。

正副両サイトの第1の記憶制御装置10（SW1），10（SW2）は、それぞれ対応する第2の記憶制御装置200（1），200（2）の論理ボリュームを取り込むためのマッピングテーブルTM1，TM2を備えている。正サイトに設置される第1の記憶制御装置10（SW1）は、例えば、図16（a）に示すようなマッピングテーブルTM1を備えている。副サイトに設置される第1の記憶制御装置10（SW2）は、例えば、図16（b）に示すようなマッピングテーブルTM2を備えている。いずれのマッピングテーブルも、既に述べた通り、対応する第2の記憶制御装置200（1），200（2）の実

L U 2 2 2 (1), 2 2 2 (2)へのアクセスポート I D、ボリューム I D、アドレスを含んでいる。

【0093】

初期コピーを行う場合、ホスト装置1は、例えば、図16(c)に示すようなコマンド(初期コピー開始リクエスト)を生成する。このコマンドは、初期コピーの開始を要求するリクエスト I Dと、コピー元の装置を特定する装置 I Dと、コピー元の論理ボリュームを特定するボリューム I Dと、コピー先の装置を特定する装置 I Dと、コピー先の論理ボリュームを特定するボリューム I Dとを含んでいる。ここで、コピー元の論理ボリュームは、正サイトの第1の記憶制御装置10(SW1)が有する仮想的L U 7 2 (1)のボリューム(α)であり、コピー先の論理ボリュームは、副サイトの第1の記憶制御装置10(SW2)が有する仮想的L U 7 2 (2)のボリューム(β)である。

【0094】

図17を参照して処理の流れを説明する。図17に示す流れ図は、正サイトの第1の記憶制御装置10(SW1)のコマンド制御部100(1)により実行される処理の概略を示している。ホスト装置1が図16(c)に示すようなコマンドを正サイトの第1の記憶制御装置10(SW1)に送信すると、コマンド制御部100(1)は、受信したコマンドを解析し、データの読み出し位置の初期設定を行う(S11)。即ち、コマンド制御部100(1)は、ホスト装置1から受信したコマンド中のコピー元装置I D及びコピー元ボリュームI Dから、正サイトの第1の記憶制御装置10(SW1)が提供する仮想的L U 7 2 (1)(ボリュームI D = α)を特定し、この仮想的L U 7 2 (1)の先頭論理ブロックアドレスをデータ読み出し位置として初期設定する。

【0095】

次に、コマンド制御部100(1)は、初期設定したデータの読み出し位置から、予め一回分の送信データ量として定められている量のデータを読み出す(S12)。ここで、コマンド制御部100(1)は、マッピングテーブルTM1を参照することにより、仮想的L U 7 2 (1)に割り当てられている実L U 2 2 2 (1)を検索し、実L U 2 2 2 (1)から1回分のデータを読み出す。

【0096】

そして、コマンド制御部100(1)は、ホスト装置1から指定されたコピー先、即ち、副サイトに設置された第1の記憶制御装置10(SW2)の仮想的L U 7 2 (2)(ボリュームI D = β)に向けて、先ほど読み出したデータを送信し、書き込ませる。即ち、通常の書込みコマンドにより、データの書込みを副サイトの第1の記憶制御装置10(SW2)に要求する。この書込みコマンドで指定されるボリュームI Dは β であり、書込みアドレスは先頭の論理ブロックアドレスである。

【0097】

正サイトの第1の記憶制御装置10(SW1)のイニシエータポート12(1)から送信された書込みコマンドは、通信ネットワーク2(1), 2(2)を経由して副サイトに到達し、副サイトの第1の記憶制御装置10(SW2)のターゲットポート11(2)に受信される。

【0098】

第1の記憶制御装置10(SW2)のコマンド制御部100(2)は、書込みコマンドを解析し、マッピングテーブルTM2を参照して、仮想的L U 7 2 (2)(ボリュームI D = β)に対応する実L U 2 2 2 (2)(ボリュームI D = β 2)を検索する。そして、コマンド制御部100(2)は、受信したデータを実L U 2 2 2 (2)に書き込ませるべく、データの書込みコマンドを生成し、書込みコマンドをイニシエータポート12(2)から第2の記憶制御装置200(2)に送信する(S13)。ターゲットポート201(2)を介して第2の記憶制御装置200(2)が書込みコマンドを受信すると、コマンド制御部210(2)は、受信したデータを実L U 2 2 2 (2)の先頭論理ブロックアドレスに記憶させる。書込みが完了すると、第2の記憶制御装置200(2)は、書込み完了報告を第1の記憶制御装置10(SW1)に送信する。また、副サイトの第1の記憶制御

装置10（SW2）は、書き込み完了報告を正サイトの第1の記憶制御装置10（SW1）に送信する。なお、書き込み完了報告の送信タイミングは、場合に応じて同期式でも非同期式でもよい。

【0099】

1回分の書き込みが完了すると、正サイトの第1の記憶制御装置10（SW1）のコマンド制御部100（1）は、前回読み出した位置から一回分のデータ量の分だけ位置を進め、データ読み出し位置を更新する（S14）。コピー元の論理ボリュームの最終論理ブロックアドレスを超えるまで、以上のS12～S14の処理を繰り返すことにより、ペアボリューム間の初期コピーが完了する。

【0100】

このように、正副両サイト間のペアボリュームの初期コピーでは、主導する第1の記憶制御装置10（SW1）は、S12～S14の処理を何度も繰り返す必要があり、その負担は大きい。そこで、次に、第1の記憶制御装置10の負荷を低減すべく、より改善された方法を提案する。

【0101】

〔論理ボリュームコピー2〕

図18～図21に基づいて、第1の記憶制御装置10の負荷を軽減してペアボリュームの初期コピーを行う場合を説明する。

【0102】

上記と同様に、正副両サイトの第1の記憶制御装置10（SW1），10（SW2）は、それぞれ対応する第2の記憶制御装置200（1），200（2）の有する実LU222（1），222（2）を自己の仮想的LU72（1），72（2）として利用するために、図19（a），（b）に示すようなマッピングテーブルTM1，TM2をそれぞれ備えている。このマッピングテーブルTM1，TM2は、記憶システムを構成する際に登録される。また、図19（c），（d）に示すように、正副両サイトの第1の記憶制御装置10（SW1），10（SW2）は、それぞれ対応する第2の記憶制御装置200（1），200（2）がサポートしている機能の一覧を管理する機能管理テーブルTF1，TF2を備えている。即ち、正副両サイトのそれをおいて、第1の記憶制御装置10（SW1），10（SW2）は、第2の記憶制御装置200（1），200（2）の実LU222（1），222（2）を間接的に支配下に置くと共に、第2の記憶制御装置200（1），200（2）の有するサポート機能を把握している。

【0103】

ホスト装置1は、上記と同様に、例えば、図19（e）に示すような構造の初期コピー開始コマンドを生成し、通信ネットワーク2（1）を介して正サイトの第1の記憶制御装置10（SW1）に送信する。この初期コピー開始コマンドが、ターゲットポート11（1）を介して第1の記憶制御装置10（SW1）に受信されると、コマンド制御部100（1）は、受信したコマンドを解析し、初期コピー処理を開始させる。以下の動作の流れについて、図20も参照しつつ説明する。図20は、初期コピー開始コマンドを受信したコマンド制御部100（1）により実行される処理の概要を示す。

【0104】

コマンド制御部100（1）は、最初に、正副両サイトに設置された第2の記憶制御装置200（1），200（2）がサポートする機能の確認を行う（S21）。コマンド制御部100（1）は、図19（c）に示す機能管理テーブルTF1を参照し、正サイト側の第2の記憶制御装置200（1）が論理ボリュームコピー機能をサポートしているか否かを確認する。例えば、機能F1を論理ボリュームコピー機能とする、「可」が設定されているため、コマンド制御部100（1）は、正サイト側の第2の記憶制御装置200（1）に論理ボリュームコピー処理を代行させることが可能であることを知る。

【0105】

次に、コマンド制御部100（1）は、イニシエータポート12（1）から通信ネットワーク2（1），2（2）を介して、機能管理テーブルTF2の取得を要求するコマンド

(テーブル取得リクエスト) を、副サイト側の第1の記憶制御装置100 (SW2) に送信する。この取得コマンドは、ターゲットポート11 (2) を介してコマンド制御部100 (2) に入力される。コマンド制御部100 (2) は、図19 (d) に示す機能管理テーブルTF2の内容を、取得コマンドに対する応答として正サイト側に返信する。正サイト側のコマンド制御部100 (1) は、副サイト側から取得した機能管理テーブルTF2の内容に基づいて、副サイト側の第2の記憶制御装置200 (2) が論理ボリュームコピー機能を備えていることを把握する。取得された機能管理テーブルTF2の内容は、第1の記憶制御装置100 (SW1) の共有メモリまたはキャッシュメモリに格納される。なお、機能管理テーブルTF2の全ての内容を副サイト側から正サイト側に送信する場合に限らず、例えば、正サイト側のコマンド制御部100 (1) から副サイト側のコマンド制御部100 (2) に対し、副サイト側の第2の記憶制御装置200 (2) が特定の機能をサポートしているか否かについて問い合わせるようにしてもよい。

【0106】

このようにして、コマンド制御部100 (1) は、正サイトと副サイトの両方に設置された第2の記憶制御装置200 (1), 200 (2) がそれぞれ論理ボリュームコピー機能を備えているか否かを確認し、両方共に論理ボリュームコピー機能を備えている場合には (S22:YES) 、以下に述べるように、第2の記憶制御装置200 (1), 200 (2) による直接的な論理ボリュームコピーを開始させる。

【0107】

コマンド制御部100 (1) は、ホスト装置1からコピー先ボリュームとして指定された副サイト側の仮想的LU72 (2) に対応する実LU222 (2) の情報を取得すべく、副サイト側の第1の記憶制御装置100 (SW2) に対して、マッピングテーブルTM2の取得を要求する (S23)。このテーブル取得リクエストは、機能管理テーブルTF2の場合と同様に行われる所以、詳細を割愛する。

【0108】

次に、コマンド制御部100 (1) は、初期コピー開始コマンドを、イニシエータポート12 (1) から通信ネットワーク2 (1) を介して第2の記憶制御装置200 (1) のターゲットポート201 (1) に送信する (S24)。この際、コマンド制御部100 (1) は、正サイト側のマッピングテーブルTM1から抽出されたコピー元論理ボリューム (α) に関する情報と、副サイト側のマッピングテーブルTM2から抽出されたコピー先論理ボリューム (β) に関する情報を合わせて、正サイト側の第2の記憶制御装置200 (1) に送信する。第2の記憶制御装置200 (1) によって初期コピーが行われている間に、ホスト装置1によって仮想的LU72 (1) の内容が更新された場合は、その更新された論理ブロックに関する情報が差分ビットマップテーブルTBに記憶される。

【0109】

正サイト側の第2の記憶制御装置200 (1) は、ターゲットポート201 (1) を介して初期コピー開始コマンドを受信すると、このコマンドはコマンド制御部210 (1) により解析され、図21に示す初期コピー処理が開始される。

【0110】

図21は、正サイト側の第2の記憶制御装置200 (1) のコマンド制御部210 (1) により実行される初期コピー処理を示す。コマンド制御部210 (1) は、初期コピー開始コマンドで指定されたコピー元ボリュームIDと、S24で取得した各論理ボリュームの情報に基づいて、コピー元である仮想的LU72 (1) (ボリュームID = α) の先頭論理ブロックアドレスをデータ読み出し位置として初期設定する (S31)。

【0111】

次に、コマンド制御部210 (1) は、設定したデータの読み出し位置から、予め一回分の送信データ量として設定されている量のデータを読み出す (S32)。このとき、コマンド制御部210 (1) は、第1の記憶制御装置100 (SW1) から先に取得したマッピングテーブルTM1の情報に基づいて、仮想的LU72 (1) に対応する実LU222 (1) を検索し、実LU222 (1) (ボリュームID = α 2) の先頭論理ブロックアド

レスから一回の送信分のデータを読み出す。

【0112】

そして、コマンド制御部210(1)は、初期コピー開始コマンドで指定されたコピー先に対して、読み出したデータを送信し、書き込ませる(S33)。この書き込みリクエストは、通常の書き込みコマンドにより行われる。このとき、コマンド制御部210(1)は、第1の記憶制御装置10(SW1)から取得したマッピングテーブルTM2の情報に基づいて、書き込みコマンド中の書き込み先を、副サイト側の仮想的LU72(2)に対応する実LU222(2)(ボリュームID=β2)に設定する。書き込み先のアドレスは、実LU222(1)からデータを読み出す際に指定したアドレスと同一である。

【0113】

第2の記憶制御装置200(1)のイニシエータポート202(1)から送信された書き込みコマンドは、通信ネットワーク2(1), 2(2)を経由して、副サイト側の第2の記憶制御装置200(2)のターゲットポート201(2)に受信される。第2の記憶制御装置200(2)のコマンド制御部210(2)は、受信した書き込みコマンドに従って、受信データを実LU222(2)の所定の位置に書き込み、書き込み完了を正サイト側の第2の記憶制御装置200(1)に報告する。

【0114】

一つの書き込みコマンドの処理を完了すると、コマンド制御部210(1)は、前回の読み出し位置から一回分の送信データ量の分だけ位置を進めることにより、データ読み出し位置を更新する(S34)。コマンド制御部210(1)は、実LU222(1)の記憶内容を実LU222(2)に全て移し替えるまで、S32～S34の処理を繰り返し実行する(S35)。初期コピーが完了すると、コマンド制御部210(1)は、第1の記憶制御装置10(SW1)に初期コピーの完了を報告する。

【0115】

初期コピーが完了した場合は図20に戻り、第1の記憶制御装置10(SW1)のコマンド制御部100(1)は、差分ビットマップテーブルTBを参照する。差分ビットマップテーブルTBには、初期コピー中に、ホスト装置1が本来のコピー元である仮想的LU72(1)に対して行った新規データの書き込みに関する情報、即ち、更新された論理ブロックのアドレス情報が保持されている。

【0116】

差分ビットマップテーブルTBを参照することにより、初期コピー中に仮想的LU72(1)の記憶内容が変更されたことが検出された場合、コマンド制御部100(1)は、差分ビットマップテーブルTBに示された全ての論理ブロックを、コピー先の仮想的LU72(2)に書き込ませる処理を実行する(S25)。この更新されたデータのコピーは、例えば、[論理ボリュームコピー1]で述べた方法に従って行うことができる。

【0117】

このように、初期コピー中に書き込まれた新規データを、副サイト側に書き込む処理を行うのは、次の理由による。第2の記憶制御装置200(1)中のコマンド制御部210(1)によって行われる初期コピーは、論理ブロックアドレスが小さいものから大きいものへと順々に行われる。例えば、初期コピーがほぼ完了に近づく頃になって、論理ボリュームの先頭論理ブロックアドレス付近に新規データの書き込みが行われると、コマンド制御部210(1)は、このデータを副サイト側にコピーすることができない。このようにして発生しうる未コピーの論理ブロックをなくすために、差分ビットマップテーブルTBに基づいた新規書き込みデータのコピーを行う。

【0118】

なお、図20において、正副両サイトの第2の記憶制御装置200(1), 200(2)のいずれか一方が、論理ボリュームコピー機能をサポートしていないと判断された場合は(S22:No)、[論理ボリュームコピー1]で述べたように、正サイト側の第1の記憶制御装置10(SW1)が主導して初期コピーを行う(S26)。

【0119】

このように構成される本実施例によれば、前記実施例と同様に、ホスト装置1から要求されたデータ処理（正サイトと副サイト間の論理ボリュームコピー）を第2の記憶制御装置200（1），200（2）に委ねることができ、データ処理の負荷を分散させて、第1の記憶制御装置10（SW1），10（SW2）の負荷を軽減することができる。従って、第1の記憶制御装置10（SW1）の情報処理能力を他のサービス実現のために用いることができ、記憶システムを有効に稼働させることができる。

【実施例3】

【0120】

次に、図22を参照して、本発明の第3実施例を説明する。本実施例の特徴は、第1の記憶制御装置10が複数の仮想的LU72（1），72（2）を備えており、これら複数の仮想的LU72（1），72（2）の記憶内容を同期化（ミラーリング）する場合に適用した点にある。

【0121】

第1の記憶制御装置10は、例えば、インテリジェント化されたファイバチャネルスイッチから構成され、2個の仮想的LU72（1），72（2）を有している。一方の仮想的LU72（1）はプライマリボリューム、他方の仮想的LU72（2）はセカンダリボリュームである。各仮想的LU72（1），72（2）の実体は、第2の記憶制御装置200の実LU222（1），222（2）である。

【0122】

仮想的LU72（1）の記憶内容を仮想的LU72（2）にコピーする場合は、実施例2の【論理ボリュームコピー1】または【論理ボリュームコピー2】で述べた方法と同様にして行うことができる。

【0123】

例えば、一つの方法によれば、第1の記憶制御装置10が、第2の記憶制御装置200の実LU222（1）のデータを先頭論理ブロックアドレスから所定量ずつ読み出し、読み出したデータを第2の記憶制御装置200の実LU222（2）の所定位置に書き込ませることにより、実LU222（1）の全データを実LU222（2）へコピーすることができる。

【0124】

また、別の方針によれば、第2の記憶制御装置200がミラーリング機能をサポートしているか否かを判別し、ミラーリング機能を有している場合は、第1の記憶制御装置10から第2の記憶制御装置200に向けてミラーリング開始コマンドを送信する。このミラーリング開始コマンドには、少なくともコピー元ボリュームID（ α ）とコピー先ボリュームID（ β ）とが含まれている。この際に、マッピングテーブルTMの記憶内容と一緒に送信することができる。あるいは、マッピングテーブルTMを参照することにより、コピー元ボリュームIDを α_2 に、コピー先ボリュームIDを β_2 にそれぞれ変換してミラーリング開始コマンドを送信してもよい。

【0125】

第2の記憶制御装置200は、ミラーリング開始コマンドを受信すると、実LU222（1）の先頭論理ブロックアドレスから最終論理ブロックアドレスまでの全データを、実LU222（2）に所定量ずつコピーする。なお、ミラーリング中に、ホスト装置1によりデータが更新された場合は、前記で述べたと同様に、差分ビットマップテーブルを利用して、更新されたデータを後からコピーすることができる。

【0126】

なお、本発明は、上述した各実施例に限定されない。当業者であれば、本発明の範囲内で、種々の追加や変更等を行うことができる。前記各実施例では、インテリジェント化されたファイバチャネルスイッチを中心に説明したが、本発明はこれに限らず、ディスクアレイ装置等にも広く適用できる。また、第1の記憶制御装置内に、それぞれ異なる第2の記憶制御装置の実記憶領域に対応する仮想的な記憶領域を設ける場合等にも適用することができる。

【図面の簡単な説明】

【0127】

【図1】本発明の第1実施例による記憶システムの要部を概略的に示すブロック図である。

【図2】第1の記憶制御装置をディスクアレイ装置から構成した場合のブロック図である。

【図3】図2に示すディスクアレイ装置をホスト装置から見た場合の論理的な概略構造を示す説明図である。

【図4】第1の記憶制御装置をファイバチャネルスイッチとして構成した場合のブロック図である。

【図5】図4に示すファイバチャネルスイッチをホスト装置から見た場合の論理的な概略構造を示す説明図である。

【図6】第1の記憶制御装置から第2の記憶制御装置に対してデータを読み書きする場合の一例を示す概略ブロック図である。

【図7】第1の記憶制御装置をディスクアレイ装置として構成し、1台の第2の記憶制御装置と接続する場合の概略ブロック図である。

【図8】第1の記憶制御装置をファイバチャネルスイッチとして構成し、仮想的LUを2つの実LUから構成する場合の概略ブロック図である。

【図9】(a)は第1の記憶制御装置をファイバチャネルスイッチとして構成した場合のマッピングテーブル、(b)は第1の記憶制御装置をディスクアレイ装置として構成した場合のマッピングテーブル、(c)は複数の実LUから仮想的LUを構成する場合のマッピングテーブルをそれぞれ示す。

【図10】第1の記憶制御装置が自らダイレクトバックアップを行う場合の説明図である。

【図11】仮想的LUを備えた第1の記憶制御装置がダイレクトバックアップを行う場合の説明図である。

【図12】仮想的LUを備えた第1の記憶制御装置が第2の記憶制御装置にダイレクトバックアップを代行させる場合の説明図である。

【図13】(a)は機能管理テーブル、(b)はホスト装置から第1の記憶制御装置に送信されるコマンドのデータ構造、(c)は第1の記憶制御装置から第2の記憶制御装置に送信されるコマンドのデータ構造をそれぞれ示す。

【図14】データのダイレクトバックアップを行う場合の第1の記憶制御装置における処理の概要を示すフローチャートである。

【図15】本発明の第2実施例に係る記憶システムの全体を概略的に示すブロック図である。

【図16】(a)は正サイト側のマッピングテーブル、(b)は副サイト側のマッピングテーブル、(c)は初期コピー開始コマンドのデータ構造をそれぞれ示す。

【図17】初期コピーを行う場合の第1の記憶制御装置における処理の概要を示すフローチャートである。

【図18】仮想的LUを備える第1の記憶制御装置が、第2の記憶制御装置に論理ボリュームのコピーを行わせる場合の記憶システムの全体概要を示すブロック図である。

【図19】(a)は正サイト側マッピングテーブル、(b)は副サイト側マッピングテーブル、(c)は正サイト側機能管理テーブル、(d)は副サイト側機能管理テーブル、(e)は初期コピー開始コマンドのデータ構造をそれぞれ示す。

【図20】第1の記憶制御装置により実行される初期コピー処理の概要を示すフローチャートである。

【図21】第1の記憶制御装置から依頼された初期コピーを行う第2の記憶制御装置の処理概要を示すフローチャートである。

【図22】本発明の第3実施例に係り、内部ボリュームの同期化を行う場合の構成概

要を示すブロック図である。

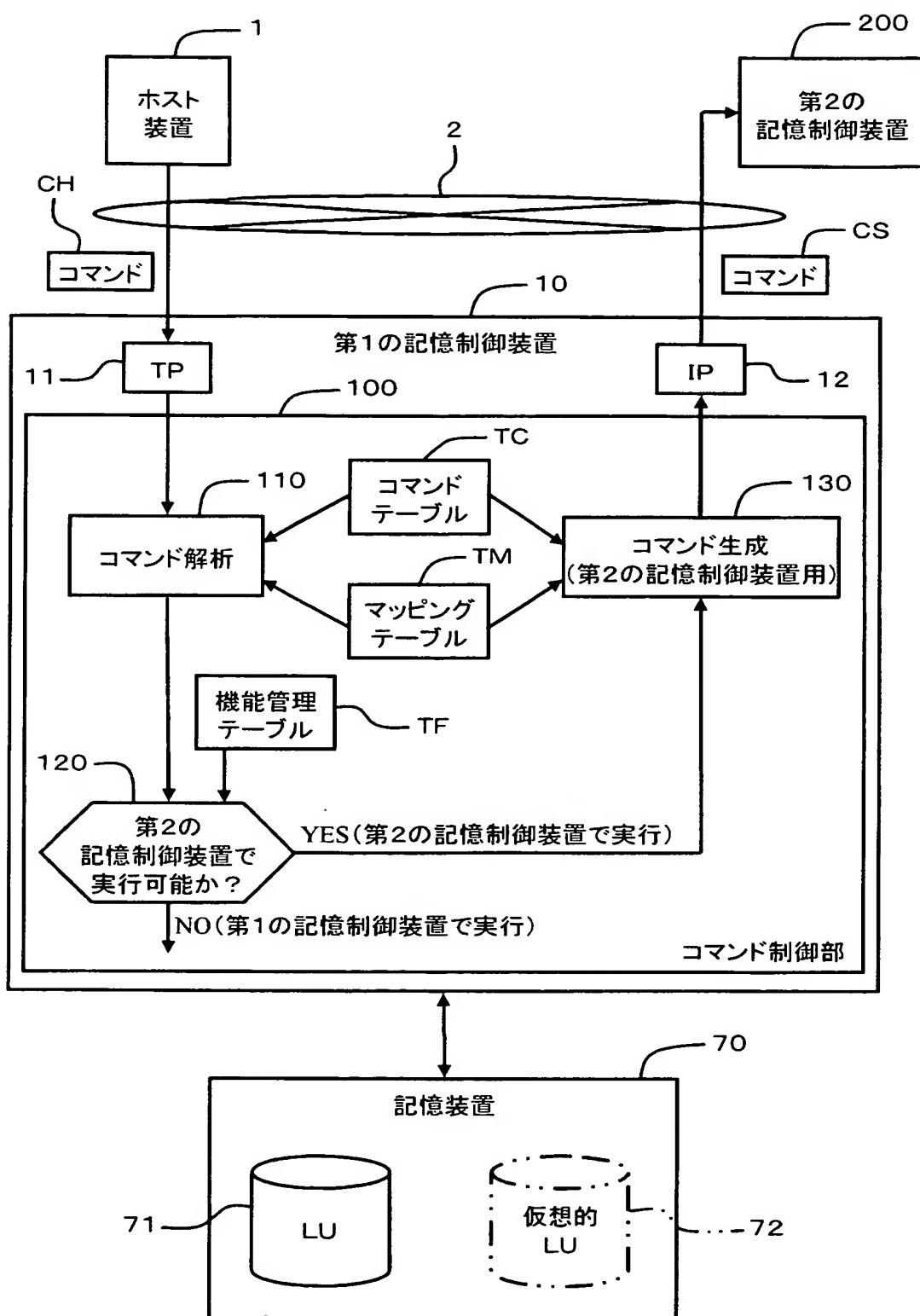
【符号の説明】

【0128】

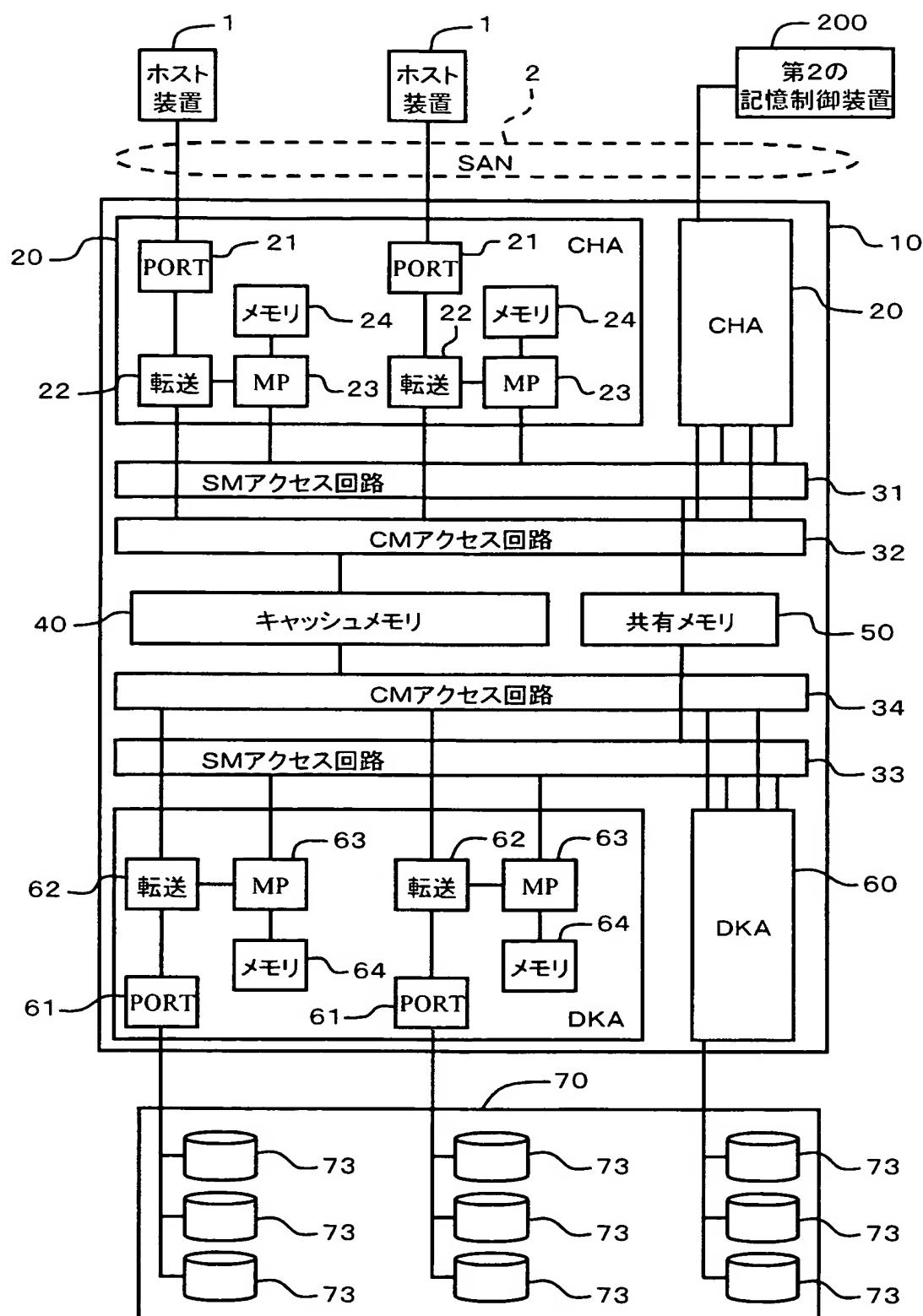
1…ホスト装置、2…通信ネットワーク、3…バックアップ装置、10…第1の記憶制御装置、20…チャネルアダプタ、40…キャッシュメモリ、50…共有メモリ、60…ディスクアダプタ、70…記憶装置、71…実論理ボリューム、72…仮想的論理ボリューム、73…ディスク装置群、80…コントローラ、100…コマンド制御部、110…コマンド解析部、120…判断部、130…コマンド生成部、200…第2の記憶制御装置、210…コマンド制御部、220…記憶装置、221…ディスク装置群、222…実論理ボリューム、TB…差分ビットマップテーブル、TC…コマンドテーブル、TM…マッピングテーブル、TF…機能管理テーブル、

【書類名】図面

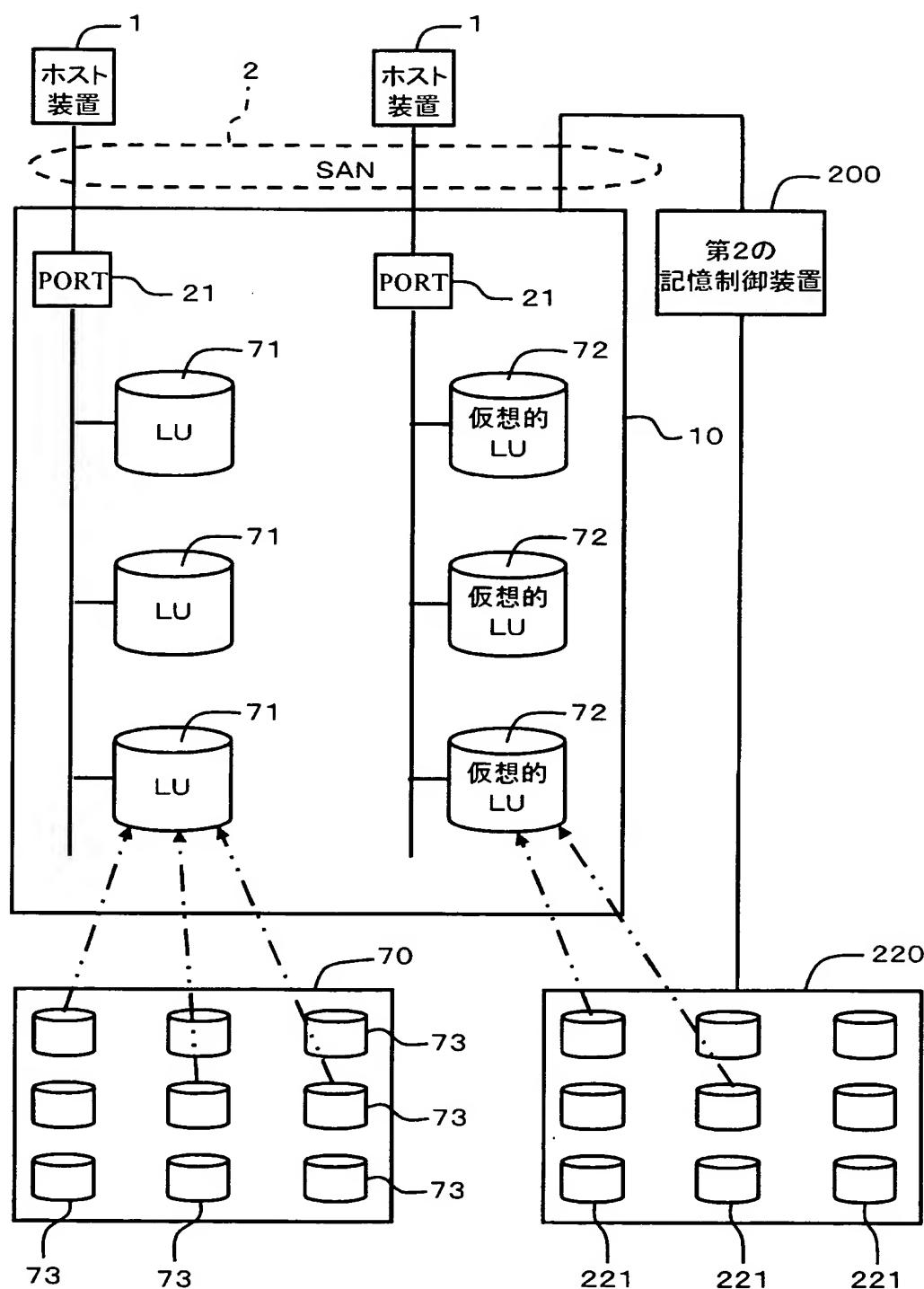
【図 1】



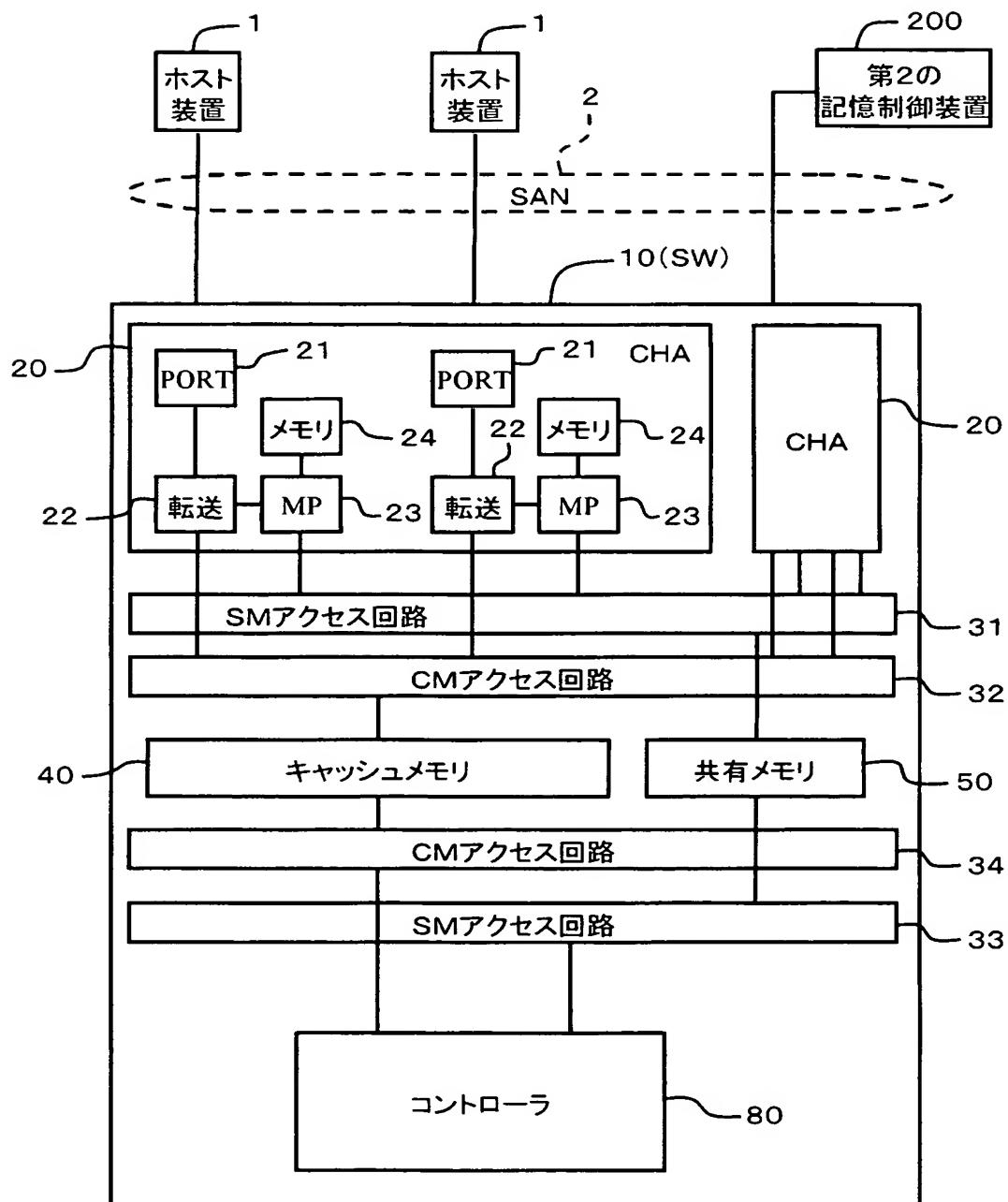
【図2】



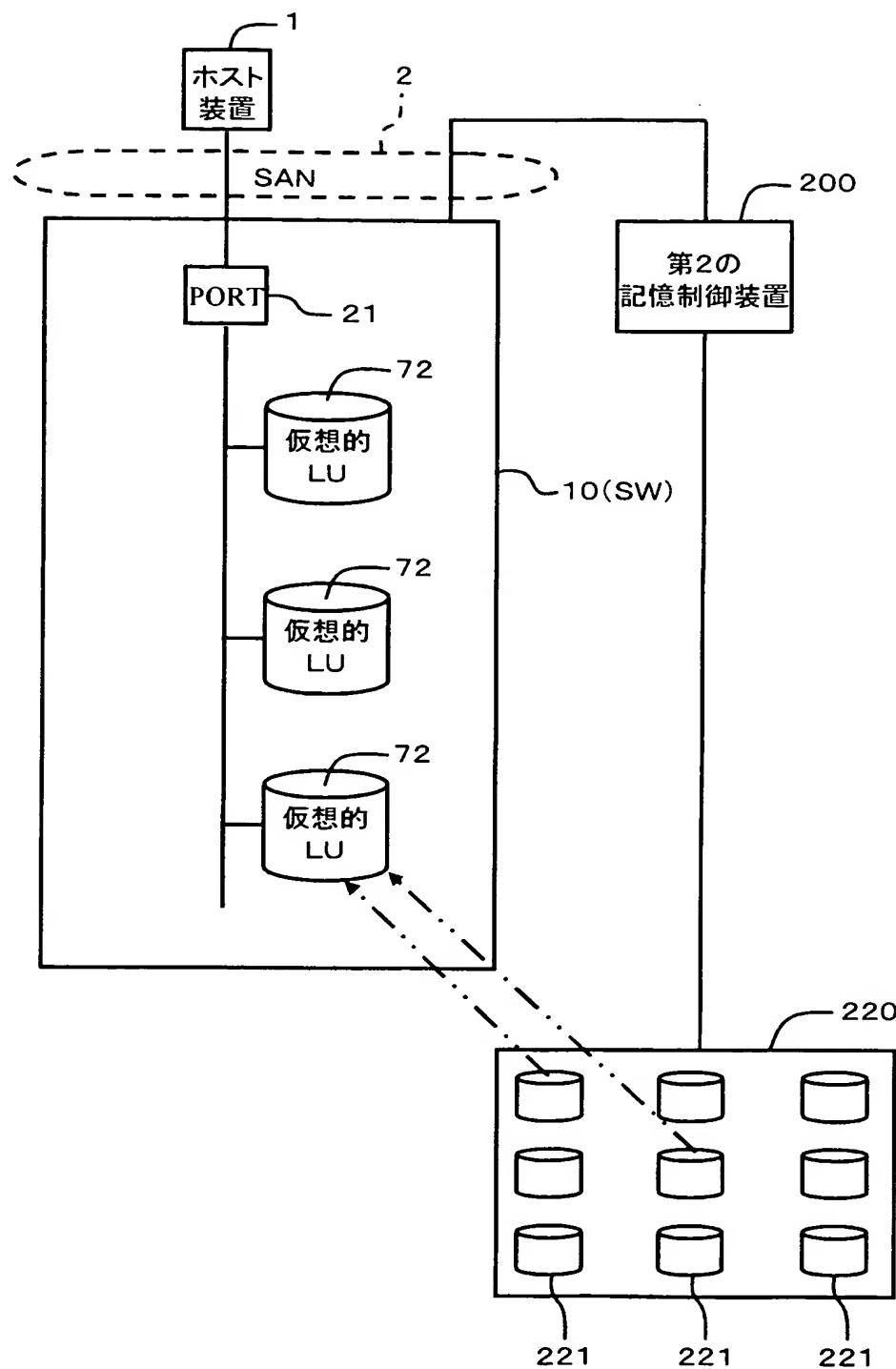
【図3】



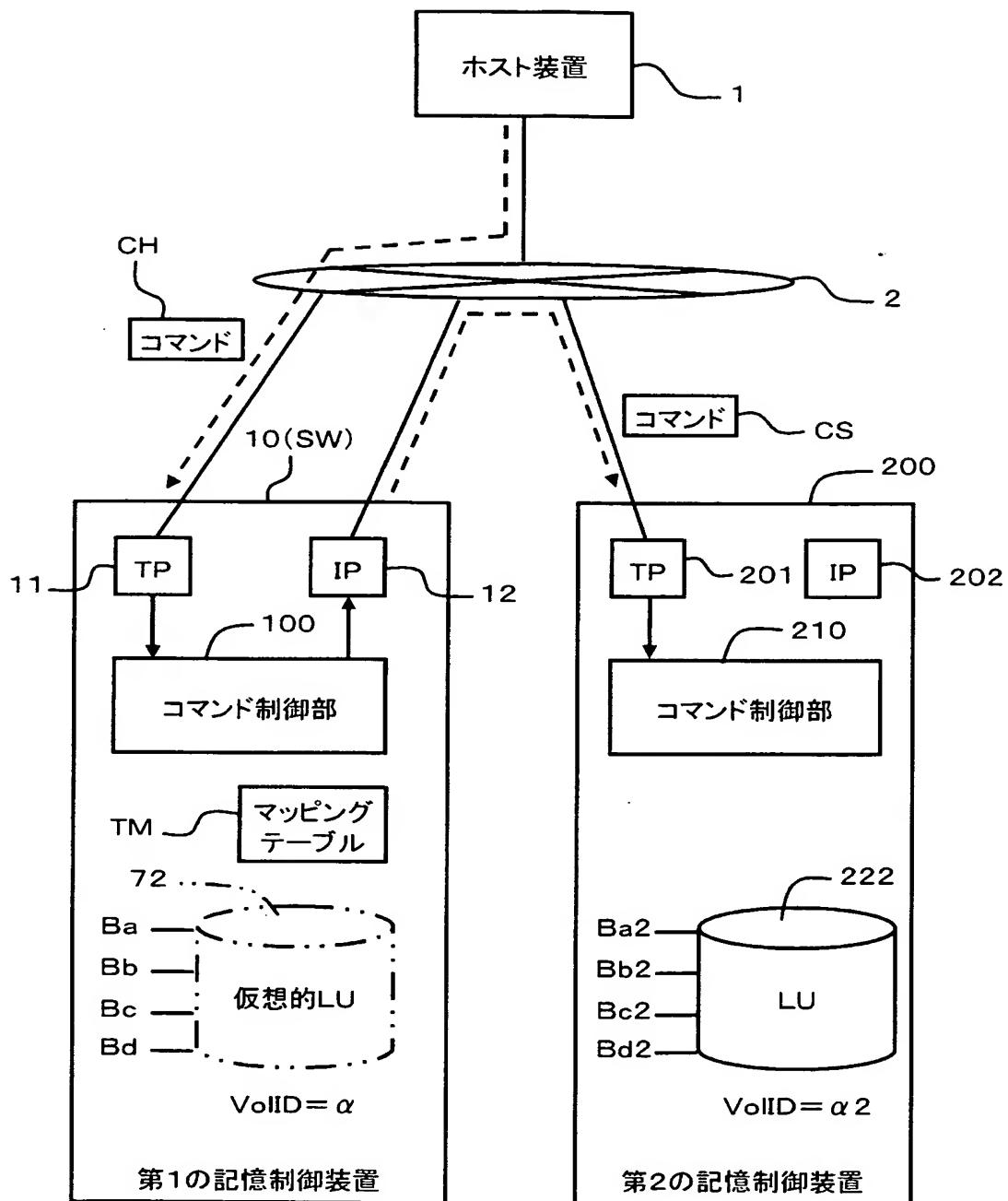
【図4】



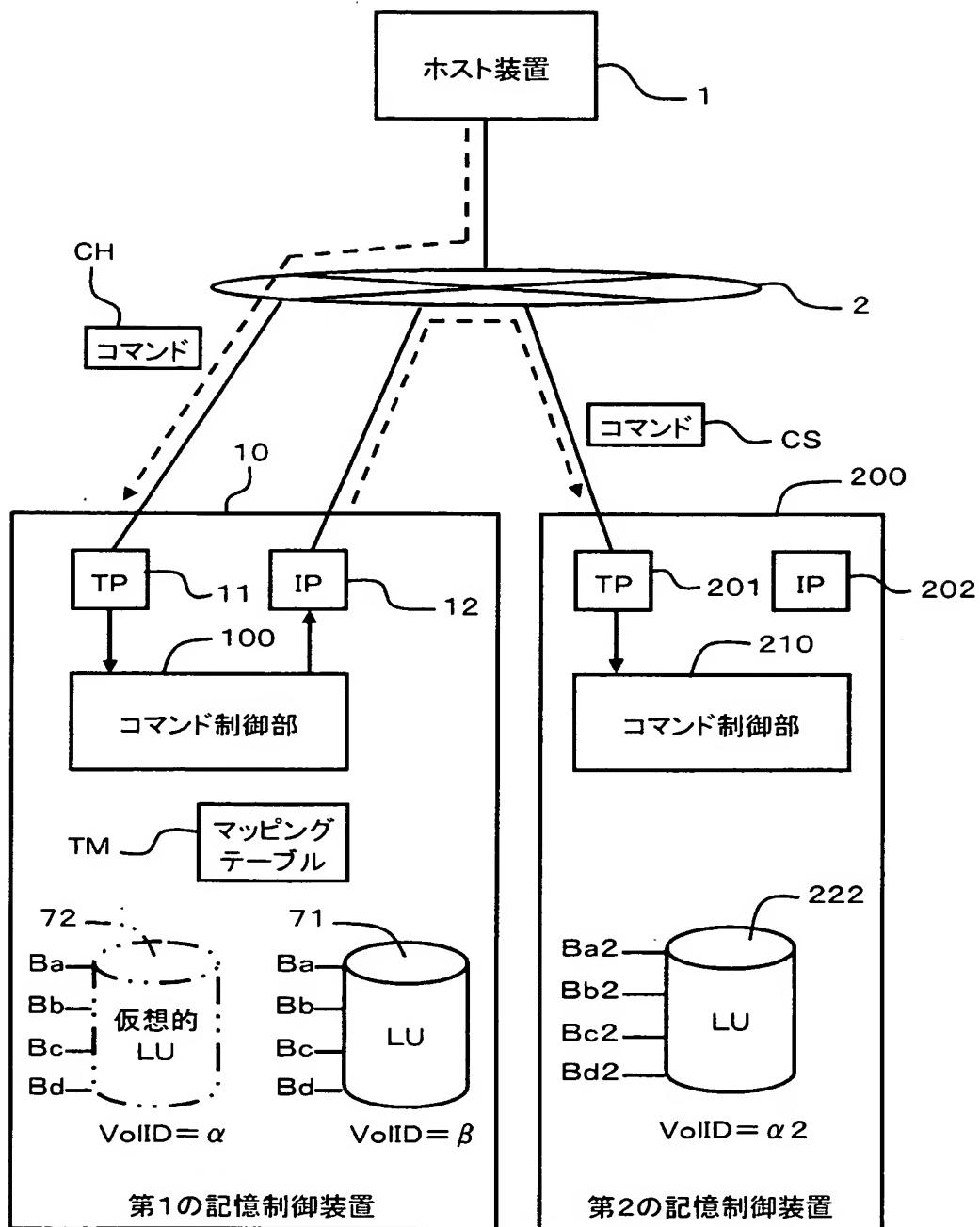
【図 5】



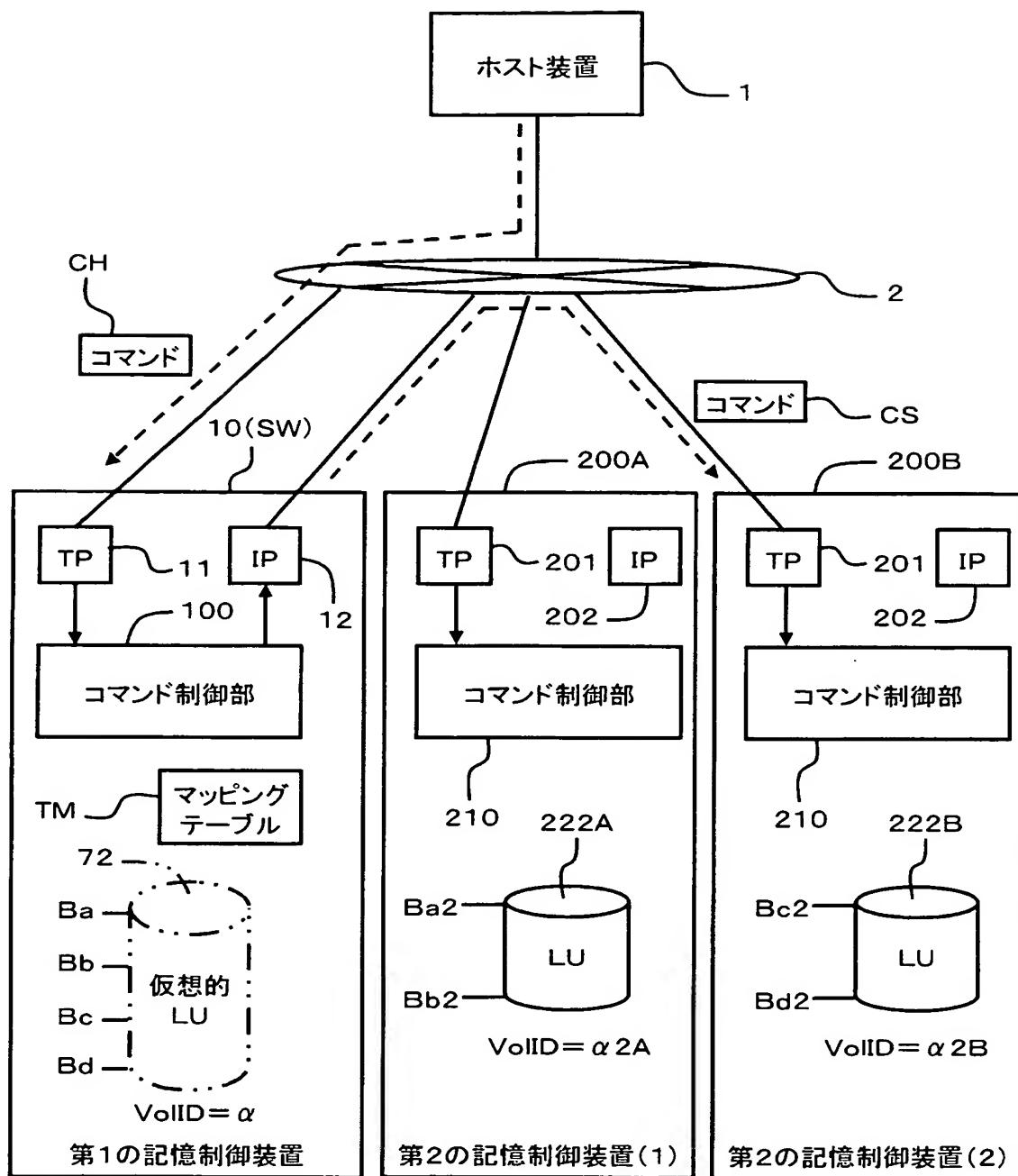
【図6】



【図7】



【図8】



【図 9】

(a)マッピングテーブルの例(ファイバチャネルスイッチの場合)

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
α	Ba	SD2	TP2	$\alpha 2$	Ba2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
α	Bd	SD2	TP2	$\alpha 2$	Bd2

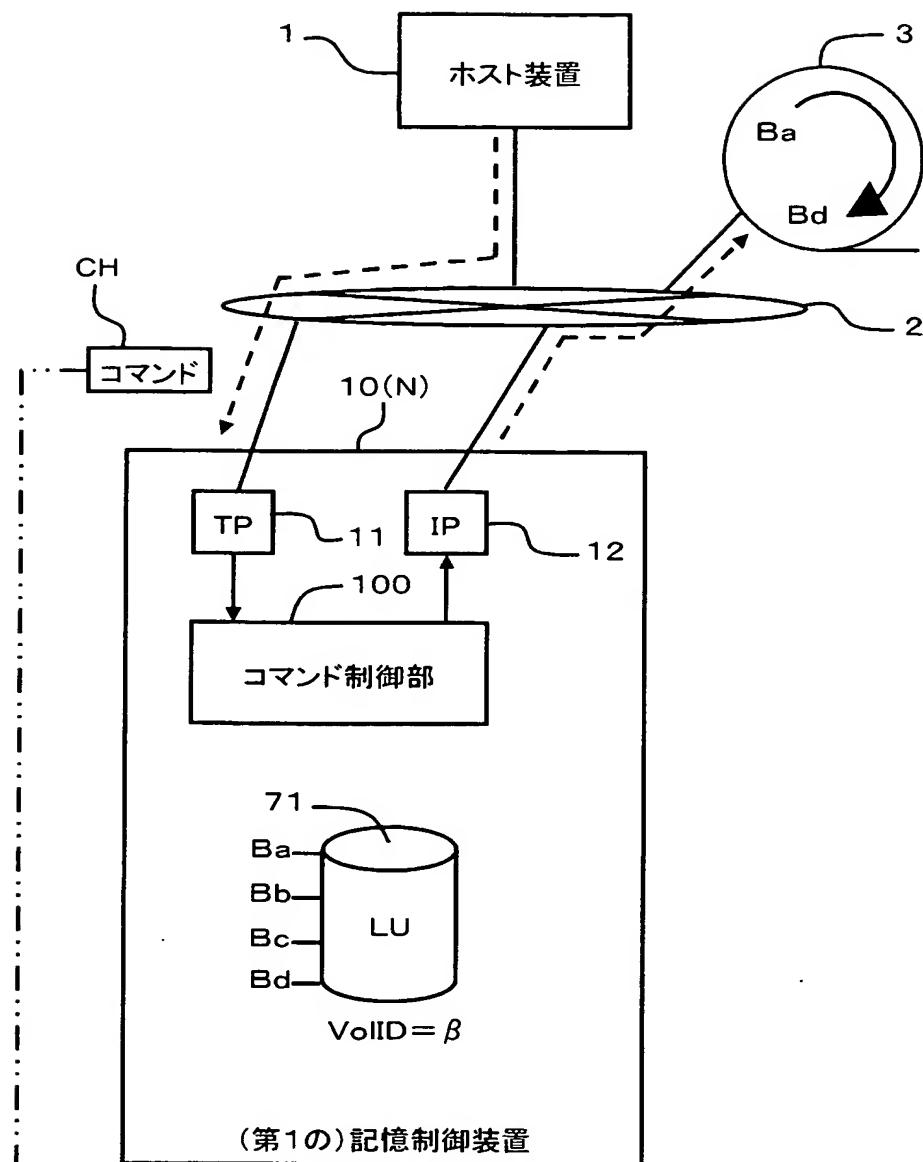
(b)マッピングテーブルの例(ディスクアレイ装置の場合)

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
α	Ba	SD2	TP2	$\alpha 2$	Ba2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
α	Bd	SD2	TP2	$\alpha 2$	Bd2
β	Ba	SD1	INTERNAL	—	—
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
β	Bd	SD1	INTERNAL	—	—

(c)マッピングテーブルの例(第2の記憶制御装置が複数ある場合)

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
α	Ba	SD2(1)	TP2(1)	$\alpha 2A$	Ba2
\vdots	Bb	SD2(1)	TP2(1)	$\alpha 2A$	Bb2
\vdots	Bc	SD2(2)	TP2(2)	$\alpha 2B$	Bc2
α	Bd	SD2(2)	TP2(2)	$\alpha 2B$	Bd2

【図10】



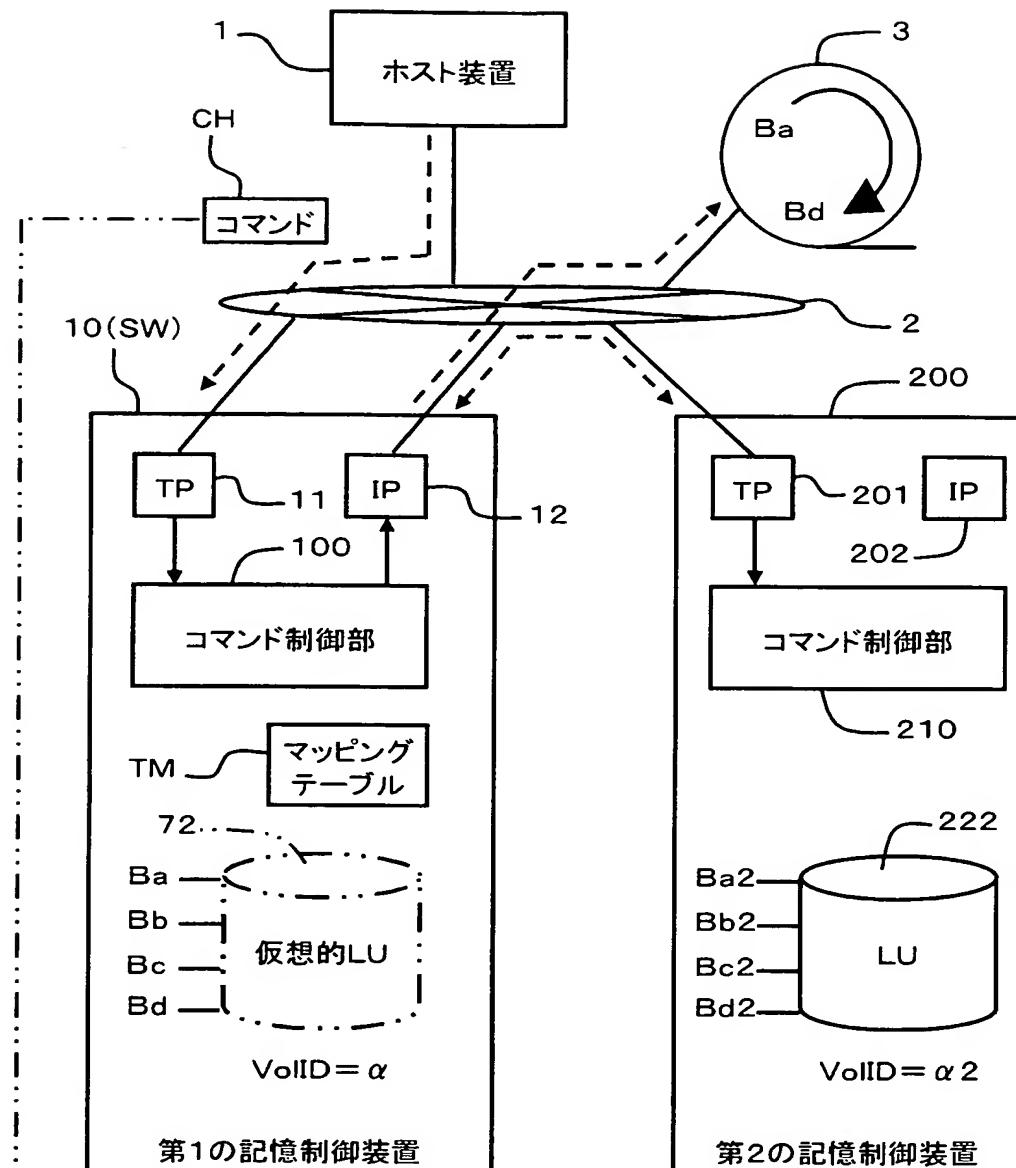
(a) データをバックアップする場合

コマンドコード (コピーリクエスト)	コピー元アドレス (Ba-Bd)	コピー先装置ID (バックアップ装置)
-----------------------	---------------------	------------------------

(b) 制御情報をバックアップデバイスに書き込む場合

コマンドコード (書き込みリクエスト)	制御情報	コピー先装置ID (バックアップ装置)
------------------------	------	------------------------

【図11】



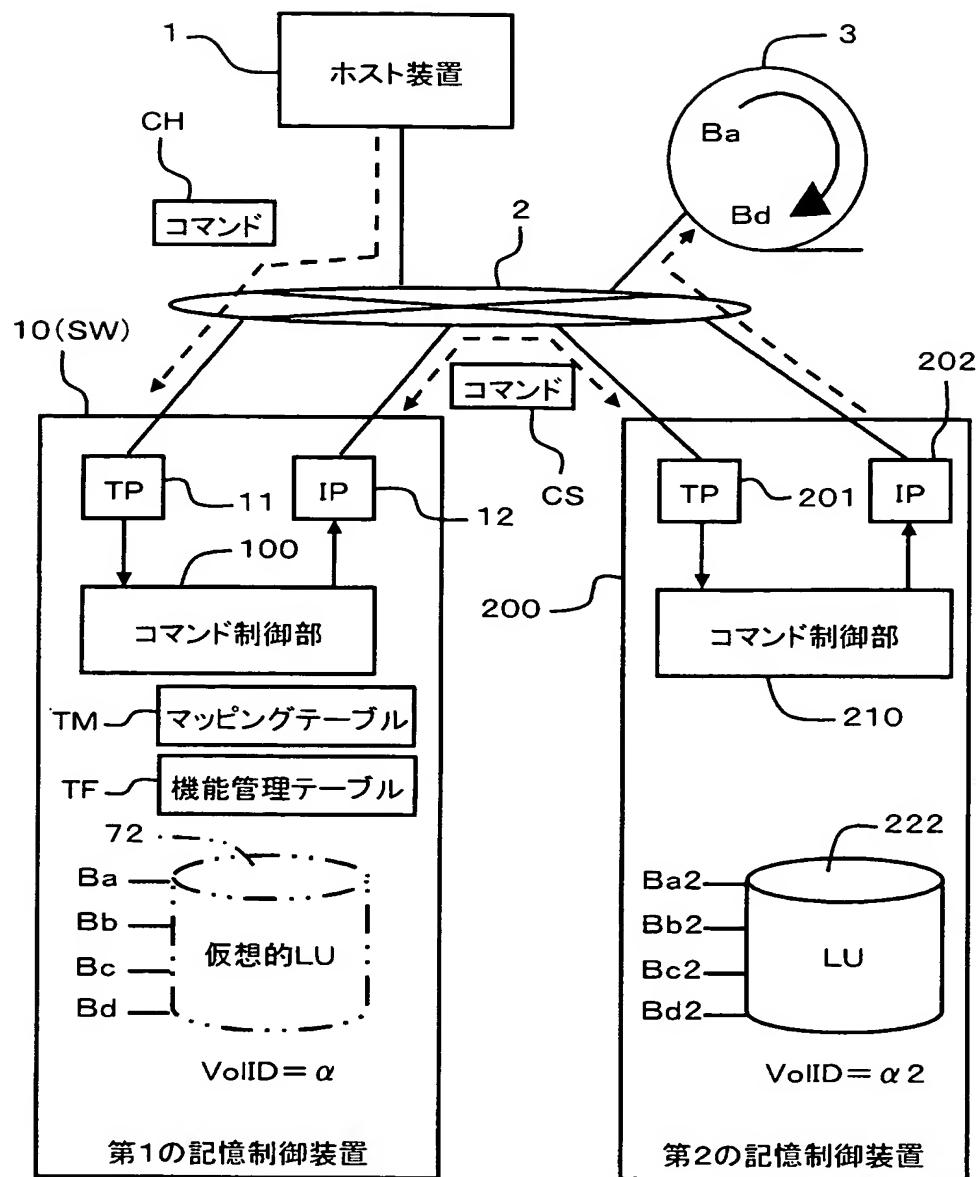
(a) データをバックアップする場合

コマンドコード (コピーリクエスト)	コピー元アドレス (Ba-Bd)	コピー先装置ID (バックアップデバイス)
-----------------------	---------------------	--------------------------

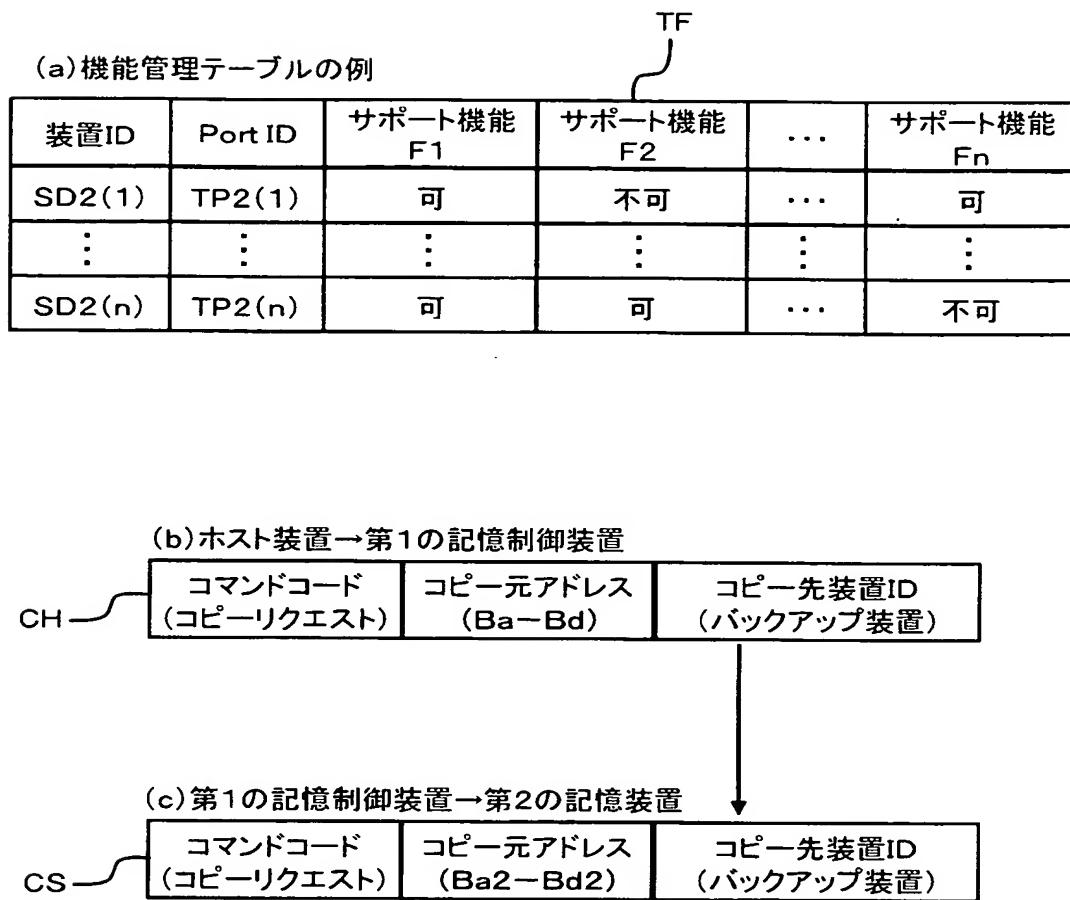
(b) 制御情報をバックアップデバイスに書き込む場合

コマンドコード (書き込みリクエスト)	制御情報	コピー先装置ID (バックアップデバイス)
------------------------	------	--------------------------

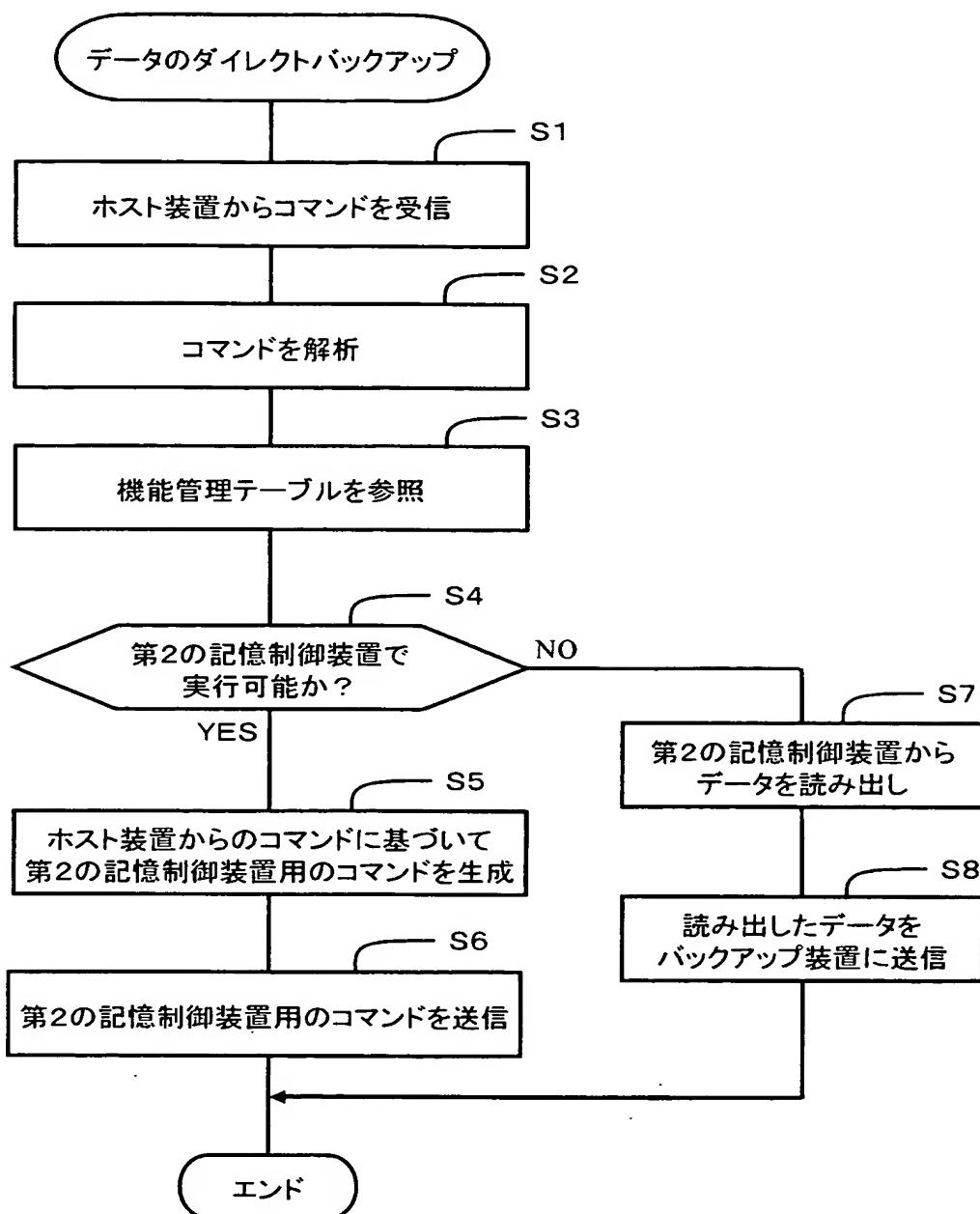
【図12】



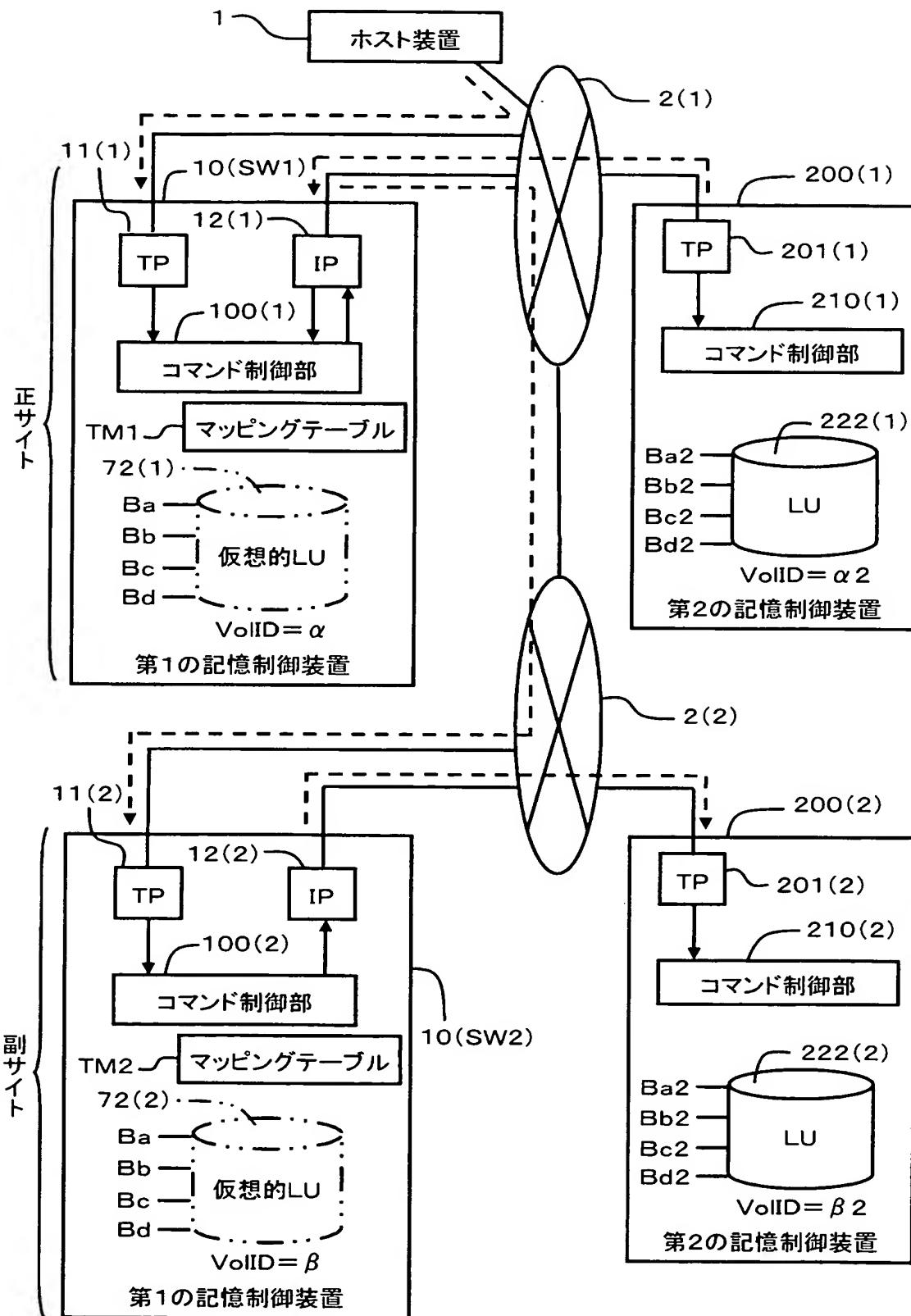
【図13】



【図 14】



【図15】



【図16】

(a) 正サイト側マッピングテーブルの例

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
α	Ba	SD2(1)	TP2(1)	$\alpha 2$	Ba2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
α	Bd	SD2(1)	TP2(1)	$\alpha 2$	Bd2

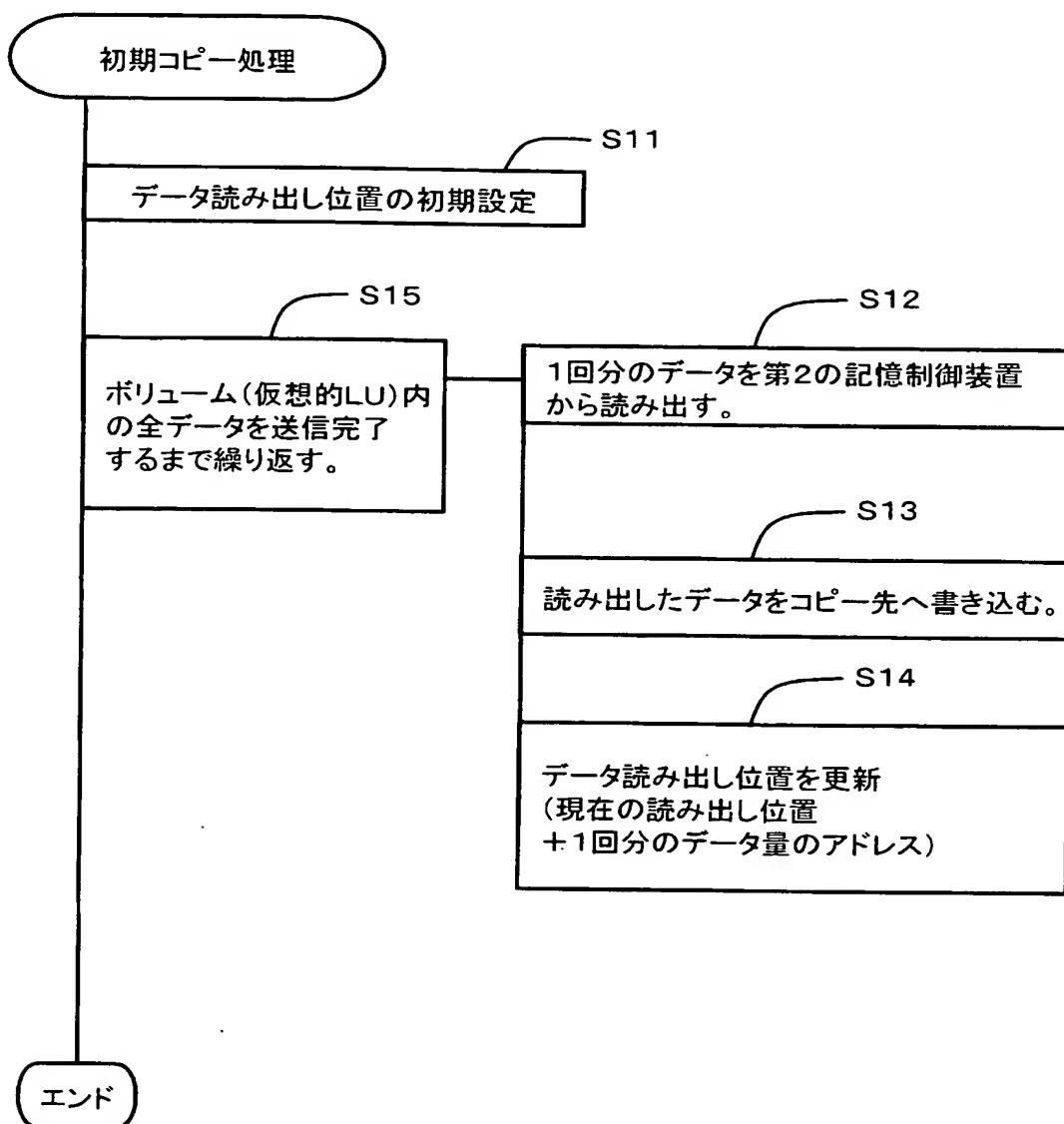
(b) 副サイト側マッピングテーブルの例

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
β	Ba	SD2(2)	TP2(2)	$\beta 2$	Ba2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
β	Bd	SD2(2)	TP2(2)	$\beta 2$	Bd2

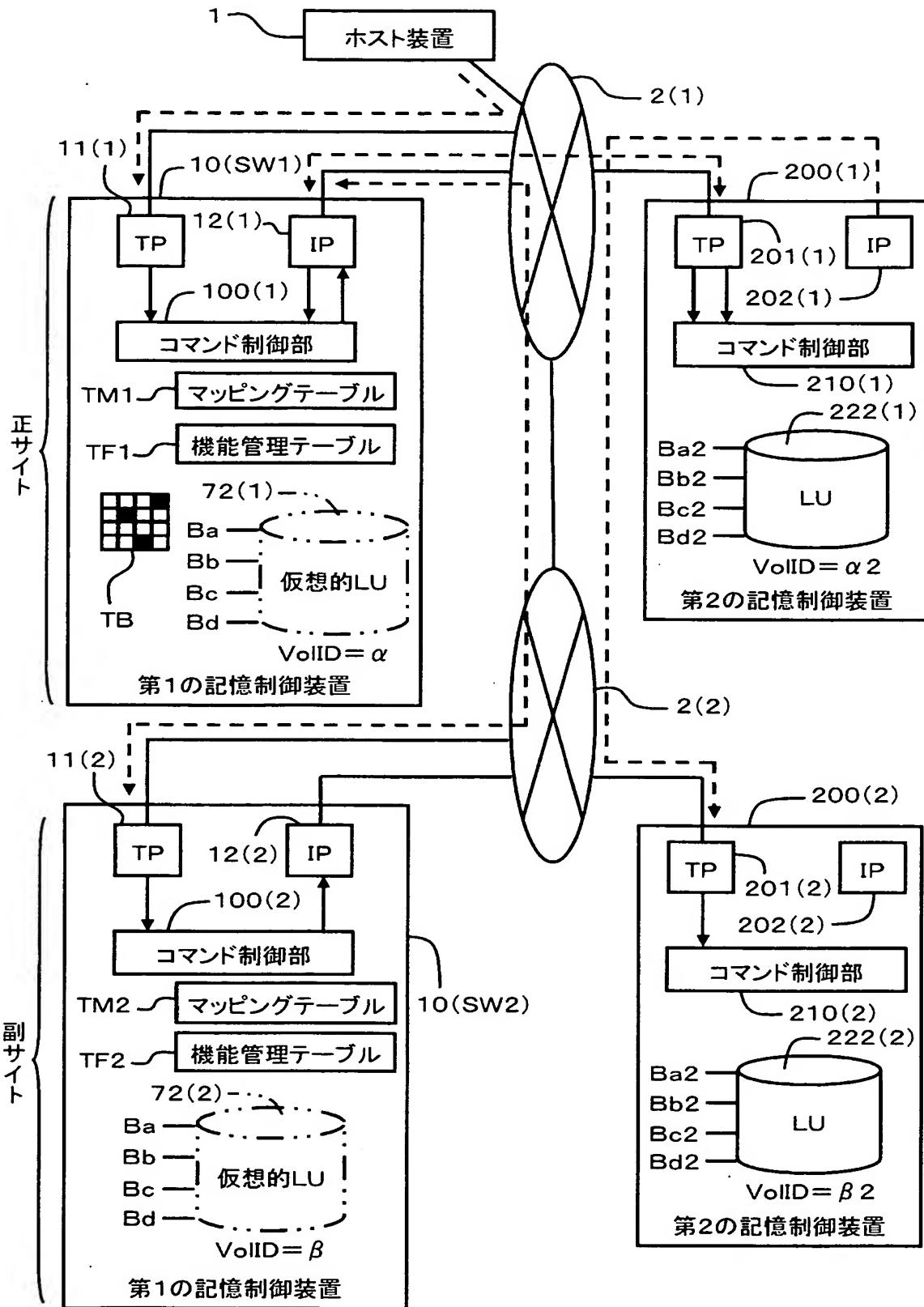
(c) 初期コピーを行う場合のコマンド構造

コマンドコード (初期コピー 開始リクエスト)	コピー元装置ID (SD1(1))	コピー元Vol ID (α)	コピー先装置ID (SD1(2))	コピー先Vol ID (β)
-------------------------------	----------------------	----------------------------	----------------------	---------------------------

【図17】



【図18】



【図19】

(a) 正サイト側マッピングテーブルの例

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
α	Ba	SD2(1)	TP2(1)	$\alpha 2$	Ba2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
α	Bd	SD2(1)	TP2(1)	$\alpha 2$	Bd2

(b) 副サイト側マッピングテーブルの例

Vol ID	BLK ADDR	装置ID	Port ID	Vol ID	BLK ADDR
β	Ba	SD2(2)	TP2(2)	$\beta 2$	Ba2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
β	Bd	SD2(2)	TP2(2)	$\beta 2$	Bd2

(c) 正サイト側機能管理テーブルの例

装置ID	Port ID	サポート機能 F1	サポート機能 F2	...	サポート機能 Fn
SD2(1)	TP2(1)	可	不可	...	可

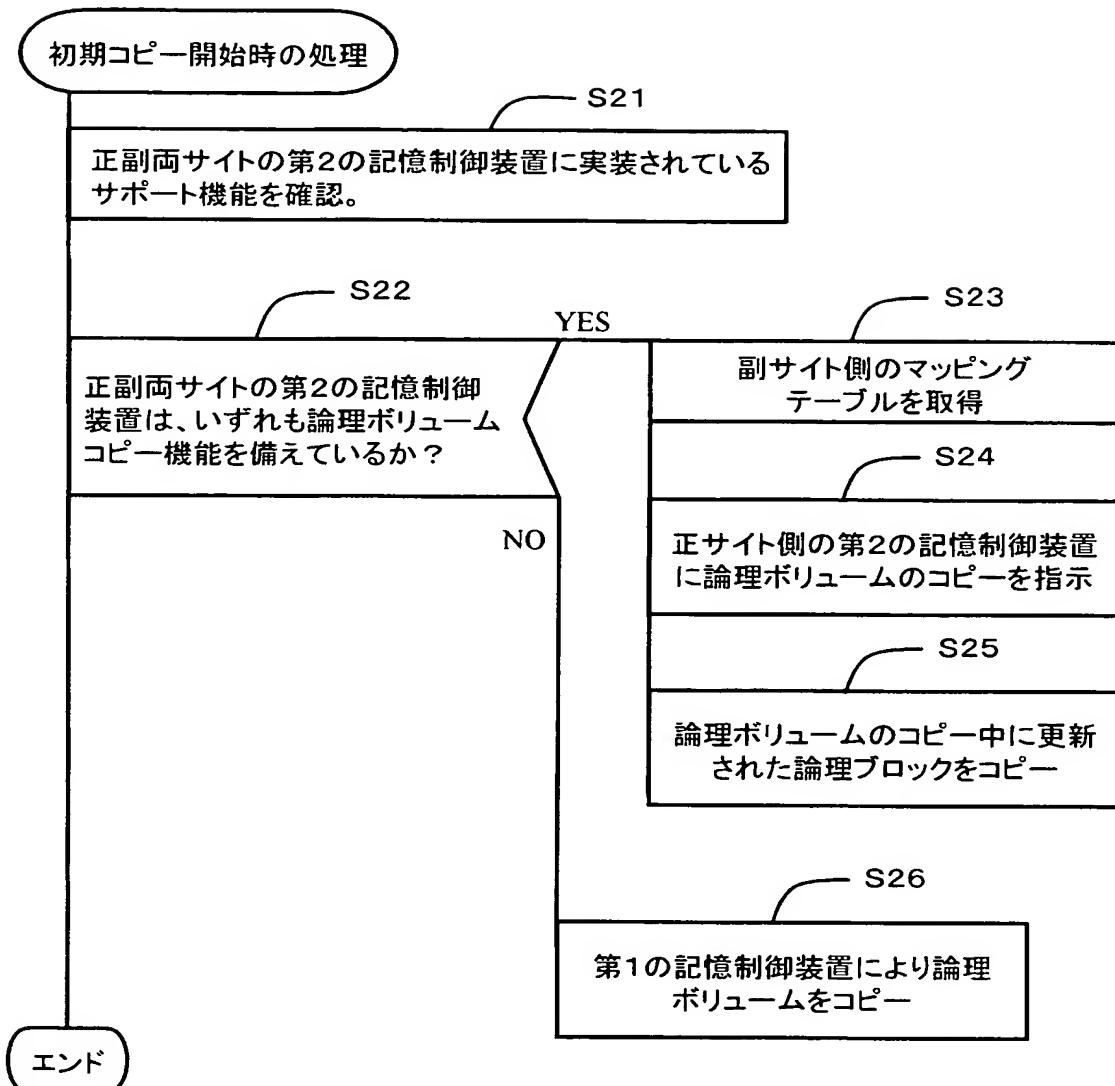
(d) 副サイト側機能管理テーブルの例

装置ID	Port ID	サポート機能 F1	サポート機能 F2	...	サポート機能 Fn
SD2(2)	TP2(2)	可	可	...	不可

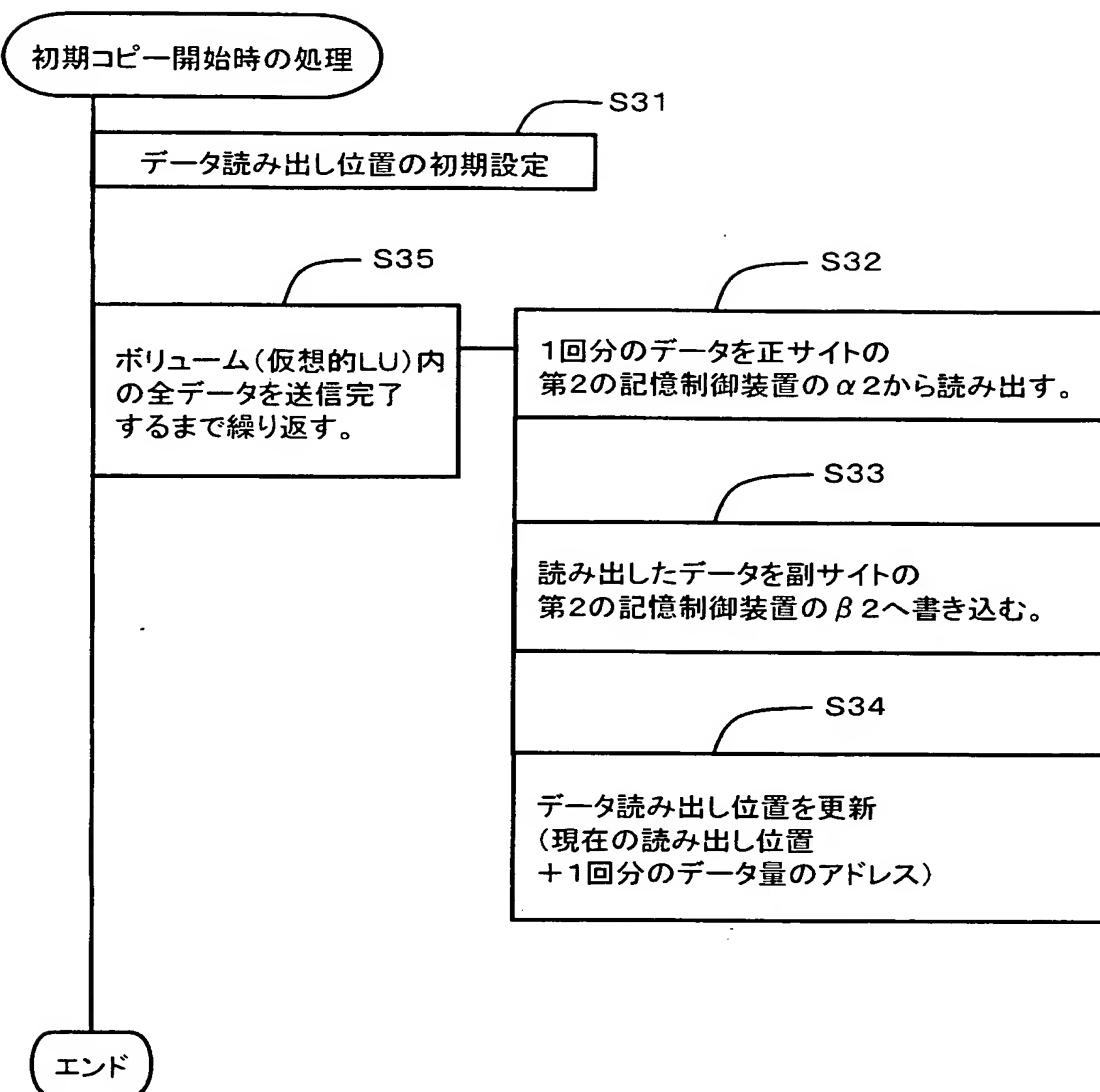
(e) 初期コピーを行う場合のコマンド構造

コマンドコード (初期コピー 開始リクエスト)	コピー元装置ID (SD1(1))	コピー元Vol ID (α)	コピー先装置ID (SD1(2))	コピー先Vol ID (β)

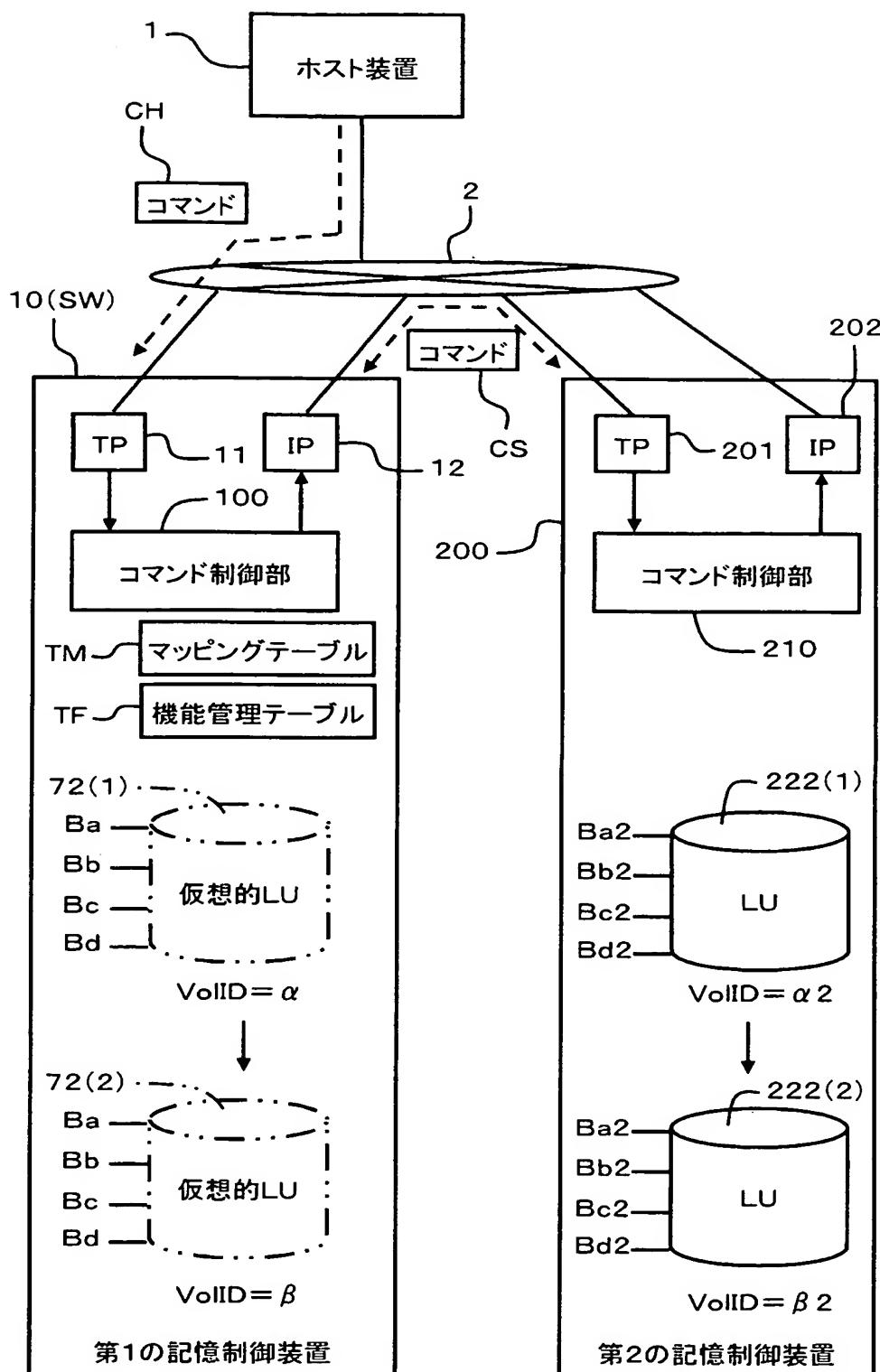
【図20】



【図21】



【図22】



【書類名】要約書

【要約】

【課題】 複数の記憶制御装置に分散された記憶領域を論理的にまとめて仮想的に提供し、特定の処理が一つの記憶制御装置に集中しないようにデータ処理を分散させる。

【解決手段】 第1の記憶制御装置10は、第2の記憶制御装置200の有する論理ボリュームを自己のボリュームであるかのように見せかける仮想的LU72を備える。ホスト装置1から要求されたデータ処理が、例えば負荷の大きな特定の処理（ダイレクトバックアップ、論理ボリュームコピー等）である場合は、機能管理テーブルTFに基づいて、要求された処理を第2の記憶制御装置で実行可能か否かを判断する。実行可能な場合は、第2の記憶制御装置にコマンドを送信し、処理を代行させる。

【選択図】 図1

認定・付加情報

特許出願の番号	特願2003-323120
受付番号	50301528271
書類名	特許願
担当官	第七担当上席 0096
作成日	平成15年 9月17日

<認定情報・付加情報>

【提出日】	平成15年 9月16日
-------	-------------

特願 2003-323120

出願人履歴情報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所